Investigating the molecular causes of inherited thrombocytopenia using CRISPR-Cas9 gene editing and super-resolution microscopy

Thesis

Abdullah O. Khan

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

Institute of Cardiovascular Science
College of Medical and Dental Sciences
University of Birmingham

Abstract

Investigating the molecular causes of inherited thrombocytopenia using CRISPR-Cas9 gene editing and super-resolution microscopy

Thesis

Abdullah O. Khan

Inherited thrombocytopenias (ITs) are a heterogenous group of disorders characterised by an abnormal platelet count (<150 x 10^9/L). Functional investigations into the mechanistic causes of ITs often suffer from numerous technical limitations, including a lack of patient material and the absence of platelet producing cell lines. In this thesis, protocols to address these issues through the application of CRISPR (Clustered Regularly Interspaced Palindromic Repeats) genome editing and stem cell differentiation are developed and applied to the study of the platelet and megakaryocyte specific β-tubulin isoform encoded by the *TubB1* gene. First, a novel method of single molecule super-resolution microscopy through the CRISPR mediated editing of cell lines is devised, and attempts to transfer this approach to stem cells are currently underway. Secondly existing methods for the differentiation of induced pluripotent stem cells into megakaryocytes are adapted to generate mature, proplatelet forming cells for *in vitro* phenotypic screening. Finally, two families from the UK GAPP (Genotyping and Phenotyping of Platelets) study with *TubB1* mutations affecting C-terminal function are investigated to determine the role of C-terminal post-translational modifications (PTMs) on ITs. Two novel PTMs (polyglutamylation and polyglycylation) which regulate motor protein processivity in both platelet production and function are reported.

Acknowledgments

Hello, dear reader. There are many people to thank, and brevity is not my friend. First and foremost, I would like to acknowledge my other half Becky. It is not easy to be partner to a one legged workaholic with a mildly deranged sense of humour, but through a religious process of self medication, she somehow manages. You have been wonderfully supportive at least 30% of the time, which I think is more than I had any right to ask for. The rest of the time, you are of course a huge pain in my backside. I would like to thank my supervisors Steve and Neil for the opportunity to do this PhD, and the freedom to do it the way I wanted to. We admittedly don't have the most conventional supervisor-student relationship... I am exceptionally lucky to have worked in a lab where we want for nothing, and have fantastic technicians who maintain our facilities and keep us well stocked with reagents, buffers, and other labware. Steph, Beata, Ying and Lourdes manage a lab of upwards of 30 people with remarkable patience - so thank you very much and my sincere apologies for being a sloppy mess. Similarly my thanks to Steve Watson for supplying and maintaining the lab so that we are so well set up to do what we do. I am also extremely fortunate to have worked with many fun and wonderful people who have now become close friends. So in no particular order, my thanks to: Natalie Poulter who is both an excellent scientist and an entertaining (if pale) human. Vicky 'Quirkle' Simms for board game nights, lunch breaks, and 7 courgettes and a handbag. Michelle Geer who was an excellent climbing buddy and one of the only other people fool enough to work here practically every weekend. Kab, Ju D'Alceste and Pete who were both excellent role models and without whom I wouldn't have been able to do the mountains of cloning I have done in the past three years (viva le Gibson!). Chiara for being Chiara, Chris for his sunny disposition and keen insights into meg biology, and Jasmeet for tea and entertainment. Alex aka Grandpere Saucisson has been a great gym buddy (in that he moves my weights around for me) and a huge help in the lab. Pip who has broken the stereotype of clinicians in the lab and who is an excellent drummer and composer for our musical project (a band temporarily named Thrombus Formation). I should also tip my hat

to my other band mates Alessandro, Iain, and Matt. Jeremy and Gayle for being as fond of extreme heavy metal as I am, as well as for your imaging expertise and ability to solve any problem respectively. From my master's project and brief time as a technician I would like to thank Josh Rappoport who encouraged independence from the very start and who also taught me that in the right lab your work mates quite quickly become good friends. Julie Mazzolini who was the best post doc you could ask for, and has since been an even better friend. Tom, Luke, and Claudio - thank you for probably the one distraction I have left to me, the cold blooded murder of virtual zombies and avatars operated by French children on Fortnite and in Blood Bowl. I suspect the list can go on indefinitely, so I will stop here and end with a thank you to my parents, sisters, and grandparents, both of whom instilled in me the value of hard work. I would like to dedicate this thesis to my grandfather, I hope one day to be worthy of your unwavering faith in me.

I believe the convention is to begin a thesis with a quote, so here are a few particularly meaningful ones:

"My centre is giving way, my right is retreating. Situation excellent, I am attacking." - *General Ferdinand Foch*

"All I know is what I read in the papers!" - *Walt Whitman*

"I want to be the very best, that no one ever was. To catch them is my real test, to train them is my cause." - *Pokemon Theme Song*

Table 1: List of Abbreviations.

ADP	Adenosine Dipshosphate
ATP	Adenosine Triphosphate
BMEC	Bone Marrow Endothelial Cell
BMP4	Bone Morphogenic Protein-4
cAMP	cyclic Adenosine Monophosphate
Cas	CRISPR associated protein
CD	Cluster of Differentiation
CLOUD	Continuum of Low Primed Undifferentiated
HSPCs	Haematopoeitic Stem and Progenitor Cells
CLP	Common Lymphoid Progenitor
CMP	Common Myeloid Progenitor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
DC	Dendritic Cells
DMEM	Dulbeccos Modified Eagles Media
DNA	Deoxyribonucleic acid
DSB	Double Strand Break
dSTORM	direct Stochastic Optical Reconstruction microscopy
EPO	Erythropoeitin
ExSTED	Expansion Stimulated Emission and Depletion Microscopy
FBS	Foetal Bovine Serum
FWHM	Full Width Half Maximum
FGF	Fibroblast Growth Factor
Flt-3	Fms Related Tyrosine Kinase-3
GAPP	Genotyping and Phenotyping of Platelets

GMP	Granulocyte Macrophage Progenitor
GMP	Guanosine Monophosphate
GPCRs	G-Protein Coupled Receptors
GPVI	Glycoprotein VI
GTP	Guanosine Triphosphate
HDR	Homology Directed Repair
HSC	Haematopoeitic Stem Cells
IL	Interleukin
IMS	Invaginated Membrane System
IPF	Immature Platelet Fraction
iPSC	induced Pluripotent Stem Cell
KI	Knock-In
KO	Knock-Out
LMPP	Lymphoid Primed Multipotent Progenitor
MEP	Megakaryocyte-Erythrocyte Progenitor
MK	Megakaryocyte
MMEJ	Micro-homology Mediated End Joining
MPP	Multipotent Progenitor
MPV	Mean Platelet Volume
MTSB	Microtubule Stabilising Buffer
NFE2	Nuclear Factor Erythroid 2
NHEJ	Non-Homologous End Joining
NK	Natural Killer Cell
NO	Nitrous Oxide
PALM	PhotoActivated Localisation Microscopy
PAM	Protospacer Adjacent Motif

PBS	Phosphate Buffered Saline
PGI3	Prostaglandin I3
PKA	Protein Kinase A
PKG	Protein Kinase G
PTM	Post-Translational Modification
RESOLFT	Reversible Saturable Optical Fluorescence Transitions
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute
RUNX	Runt-related transcription factor 1
SCF	Stem Cell Factor
SDF-1	Stromal Derived Factor-1
SDS	Sodium Dodecyle Sulphate
crRNA	CRISPR RNA
sgRNA	single guide RNA
SIM	Structured Illumination Microscopy
SMLM	Single Molecule Localisation Microscopy
SNP	Single Nucleotide Polymorphism
SOFI	Super-Resolution Fluctuation Imaging
SPIM	Selected Plane Illumination Microscopy
SRRF	Super Resolution Radial Fluctuations
ssODN	single stranded Oligo Donor Nucleotide
STED	Stimulated Emission and Depletion Microscopy
TAL	Transcription Activator Like
TALEN	Transcription Activator Like Effector Nuclease
TF	Transcription Factors
TPO	Thrombopoeitin

TxA2	Thromboxane A_2
VCAM-1	Vascular Cell Adhesion Molecule
VEGF	Vascular Endothelial Growth Factor
VLA-4	Very Late Antigen-4
vWF	von Willebrand Factor
WT	Wild Type

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Platelets are small, anucleated cells which play a vital role in haemostasis amongst other key physiological roles [1, 2]. At the onset of vascular injury, platelets play a central part in a cascade of biochemical and cellular signals and interactions - ultimately leading to the formation of a stable platelet aggregate which stems bleeding and allows for wound healing [3]. More recently, platelets have been shown to have extensive roles in diverse biological processes including inflammation and cancer metastasis[4, 5, 6]. The study of platelets has classically been limited by the absence of robust cell lines expressing megakaryocyte (MK) and platelet specific genes, and thus manipulable *in vitro* models with which to interrogate the unique functions of these cells.

This thesis is therefore focussed on the development of robust *in vitro* protocols for the study of genes and mutations of interest. The recent development of CRISPR-Cas9 offers a means by which to introduce highly specific genetic changes in an efficient and cost effective manner. Alongside recent advances in the generation of platelet producing, mature MKs from induced pluripotent stem cells, these techniques offer a means by which to bypass the lack of reliable cell line models and the artefactual effects of over-expression to study novel genes and mutations. Our particular gene of interest is the MK and platelet specific *TubB1*, and the latter part of this thesis looks at the role of novel post-translational modifications on

the function of both MKs and platelets.

In this chapter platelet production and biology are introduced, as well as the development of their progenitor cells, megakaryocytes (MKs) from haematopoeitic stem cells (HSCs) in the bone marrow. Inherited thrombocytopenias are introduced, and the role of the national Genotyping and Phenotyping of Platelets (GAPP) project in identifying novel genes and mutations is discussed in the context of a need for robust *in vitro* tools.

1.1 Introduction to Platelet Biology

Typically measuring 2-5 μm in diameter and only 0.5 μm in thickness, platelets are the smallest cellular component of circulating blood[7]. Originally identified as effectors of haemostasis and thrombosis, platelets have more recently been shown to have extensive roles in inflammation, wound healing, and cancer metastasis[8, 9]. Circulating platelets typically number between 150-450 10^9/L in adult human whole blood, and circulate in a discoid, resting state that is maintained by a distinctive microtubule marginal band[7, 10]. At the onset of vascular injury, platelets play a central role in recognising, coordinating, and responding to biochemical and cellular signals and interactions which ultimately lead to the formation of a stable platelet aggregate (thrombus) which stems bleeding and allows for wound healing (Figure 1.1A)[11, 12, 13].

In an uninjured vessel, platelets are actively inhibited by the endothelium through the production of Prostacyclin (PGI2), nitric oxide (NO), and ADPase[14, 15]. Endothelial derived PGI2 and NO stimulate adenylate and guanyl cyclase activity to increase cAMP and cGMP production respectively, which in turn activates protein kinase A (PKA) and protein kinase G (PKG)[16]. PKA and PKG inhibit activation through a series of phosphorylation events[16]. Similarly Endothelial ADPase prevents the binding of ADP to the purinergic receptor P2Y12, an event which would also lead to the accumulation of calcium and platelet activation[17, 18].

Vascular injury results in the exposure of the subendothelial space which platelet surface

receptors can interact with. These interactions trigger a series of signalling events which culminate in adhesion, spreading, secretion, and aggregation. Exposed collagen (Type I and type III) binds circulating multimers of von Willebrand Factor (vWF), resulting in its elongation and the subsequent exposure of the A1 domain to which GpIbα on the platelet surface can bind (FIgure 1.1B)[19, 20]. This is particularly important in smaller blood vessels which typically feature high shear rates. This initial binding event (also referred to as rolling and initial adhesion) allows lower affinity platelet surface receptors to interact with components of the sub-endothelial extracellular matrix (ECM), most crucially collagen.

The Fc-γ-chain associated platelet receptor glycoprotein-VI (GPVI) binds collagen type I and effects activation of a tyrosine kinase signalling cascade to achieve calcium accumulation through the activation of PLC-γ2 (Figure 1.1C)[21, 22]. GPVI signalling triggers a switch of the integrins αIIbβ3 and α2β1 to high affinity states, events which are critical to firm adhesion and the subsequent morphological changes resulting in platelet spreading[23]. Binding of αIIbβ3 to fibrinogen results in a filamin and α-actinin mediated extension of filopodia and lamellipodia which both stabilise and strengthen adhesion to the ECM, while also increasing the surface area available for platelet-platelet interaction (1.1D)[23]. α2β1 then interacts with collagen contributing to platelet adhesion. These processes occur alongside a slight expansion and subsequent retraction of the microtubule marginal band (detailed further in Chapter 4).

Platelet binding, activation, and spreading ultimately leads to the secretion of platelet granule content. Platelet secretion results in the release of ADP (from dense granules), fibrinogen, factor V, and vWF (alpha granules) amongst other components, including TxA2[24]. This release of second messengers (most critically ADP and TxA2) mediates further platelet activation through GPCRs (G-protein coupled receptors), and further stimulates coagulation with both pathways potentiating the other to achieve the formation of a stable final clot by both the amplification of platelet activation and the recruitment of circulating platelets, as well as providing components critical to the coagulation pathway. Coagulation refers to two

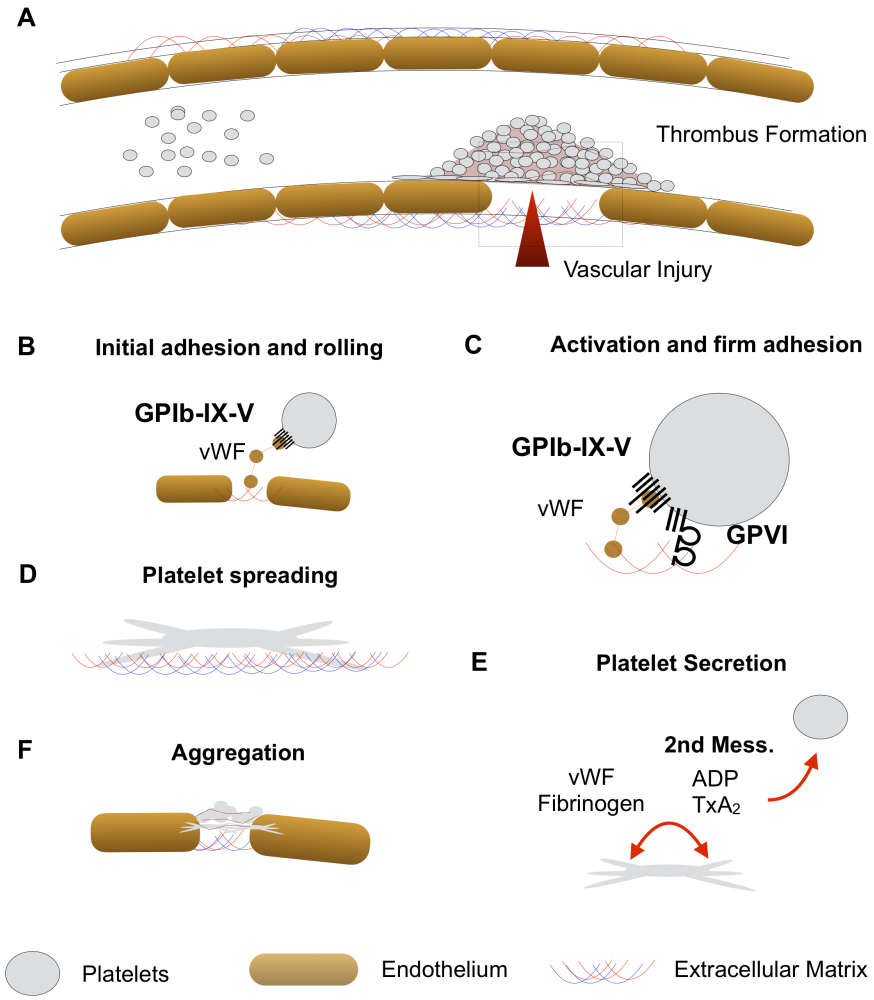

Figure 1.1: **Early stages of platelet activation and thrombus formation.**(A) Vascular injury results in the exposure of the sub-endothelial space and its constituent extracellular matrix, as well as the loss of the inhibitory effects of endothelial cells on platelet activation. (B) Exposure of collagen types I and III to circulating vWF results in its binding and subsequent elongation under shear stress. Multimers of vWF form the first point of contact with platelets through the GPIb-IX-V complex which triggers initial adhesion and rolling. (C) The subsequent exposure of platelet surface receptors to collagen results in firm adhesion through GPVI binding of collagen I. This triggers potent activation signals which ultimately result in platelet spreading. (D) Platelet spreading involves the extension of filopodia and formation of lamellopodia as the marginal band contracts for secretion. (E) Platelet secretion involves the release of the components of alpha and dense granules. These include the second messengers ADP and TxA2, both of which trigger a potent wave of signalling events involving gPCRs. Other components like vWF and fibrinogen are released to promote further platelet activation and spreading as the thrombus forms. (F) Collectively these events result in platelet aggregation and the formation of a stable thrombus, which undergoes clot retraction upon completion of the whole process.

cascades which are critical to haemostasis, the intrinsic (or contact activation) and extrinsic (tissue factor) pathways, both of which lead to the production of fibrin vital for clotting[25].

The final stages of this haemostatic process involve a localised generation of thrombin to cleave fibrinogen and secure the clot through fibrin formation. A sub-population of platelets exposed to a large influx of calcium undergo ballooning to expose phosphatidylserine (PS) on the membrane surface, causing the recruitment of coagulation factors and the cleavage of prothrombin. The increase in thrombin results in platelet activation through protease activated receptor-1 (PAR1) and its effects on G-PCRs[15, 26].

This is a unique, highly dynamic process which involves an extensive rearrangement of the platelet cytoskeleton and membrane. These morphological changes are discussed in the context of the work in this thesis in Chapter 4, and in the context of inherited thrombocytopenias in section 1.4 of this introduction. For further information, the reader is referred to an excellent review by Poulter *et al.*[27]. Once complete, a thrombus retracts through the actions of non-muscle myosin II and fibrinogen/fibrin bound αIIbβ3 to restrict clot size and minimise vessel and flow obstruction.

1.2 Megakaryopoeisis

Platelets are produced by highly specialised progenitor cells known as megakaryocytes (MKs), which are the largest and rarest of the many cells populating the bone marrow[28]. Mature MKs are typically between 50-100μM in size and make up 0.01% of all nucleated cells within the marrow[28]. In this section the development of megakaryocytes, their differentiation, and unique characteristics are discussed in the context of platelet production.

1.2.1 Megakaryocyte differentiation and development

Like most other haematopoeitic cells, megakaryocytes typically develop and mature within the bone marrow - however these cells can also migrate and as such platelet biogenesis

has been observed in peripheral blood and the lungs[29, 30, 28]. The question of the role and importance of haematopoeisis outside of the bone marrow remains a controversial one, and as such this section primarily focuses on the more established paradigm of the bone marrow as a primary site of haematopoeisis.

MKs result from the differentiation of haematopoeitic stem cells (HSCs), pluripotent, self-renewing cells which, upon the loss of their capacity for long term self-renewal, become multipotent progenitors (MPPs) which eventually give rise to the diverse subtypes of the haemotpoeitic lineage (Figure 1.2)[31]. MPPs diverge into the myeloid and lymphoid arms of the haematopoeitic lineage tree, and it is under the myeloid arm where the bi-potent megakaryocyte-erythrocyte progenitor (MEP) is found (Figure 1.2). The MEP goes on to form either erythroid unipotent or megakaryocyte progenitors.

While classically HSC lineage commitment has been thought of as a step-wise progression down a defined hierarchy or 'lineage tree', the exact nature of this model is a subject of intense debate in the field(Figure 1.2). Alternative models suggest multiple pathways can lead to the formation of MK precursors, the nature of which are also intensely debated. More recently alternative approaches suggest that specialised cells arise through a 'continuum of low primed undifferentiated haematopoeitic stem and progenitor cells' referred to as CLOUD-HSPCs(1.3)[32, 33, 31]. Ultimately the use of single cell transcriptional profiling has contradicted the paradigm of hierarchical differentiation established by 15 years of flow cytometry [31]. This model has significant implications for the mechanism by which haematological malignancies arise, as well as significant implications for the process of megakaryo- and thrombo-poeisis in the context of how the haematopoeitic niche responds to platelet deficiency and inflammation.

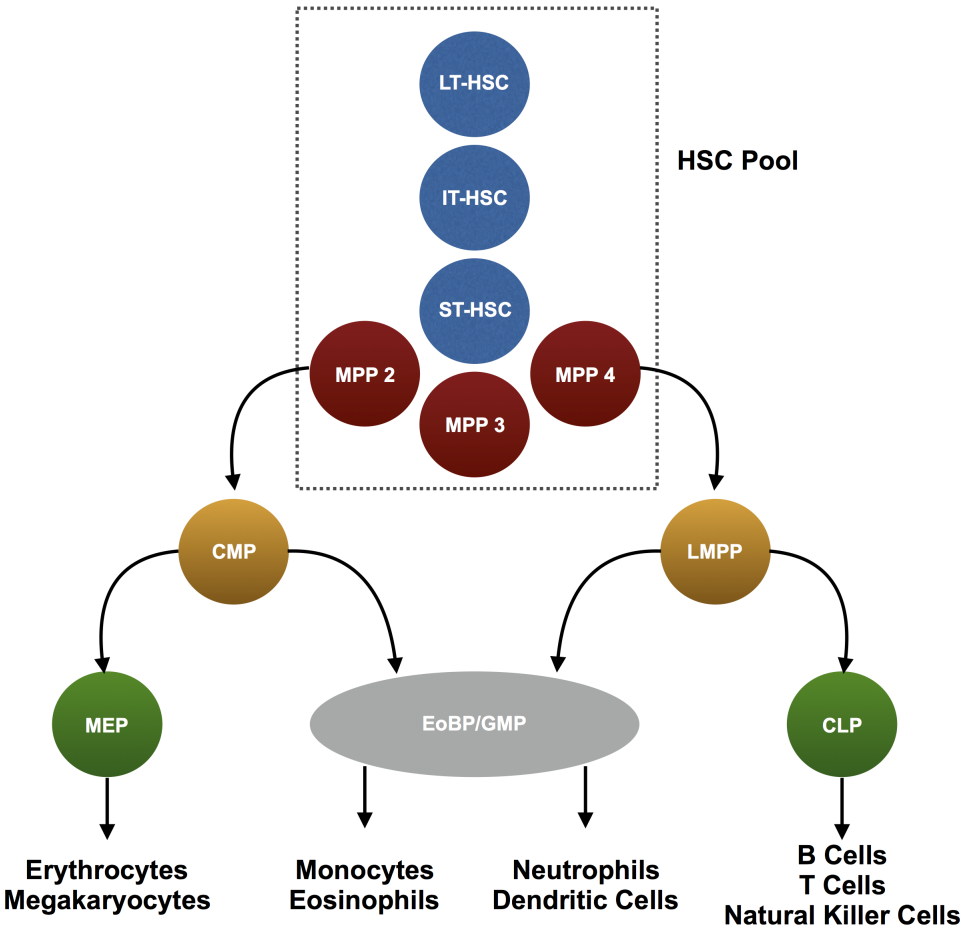

Figure 1.2: **Classic hierarchical distribution of haematopoeitic lineages.** Until very recently, the concept of haematopoeitic differentiation and lineages revolved around a hierarchical distribution whereby cells in a self-renewing, multipotent pool (HSCs of different longevities including long term (LT-HSC), intermediate (IT-HCS), and short term (ST-HSC)) commit to successively more fate limited progenitors. In later versions of this model, the HSC pool has expanded to include a range of both differentiation and self-renewal properties (demonstrated in the above tree as MPP 2, MPP 3, and MPP 4). The lymphoid primed multipotent progenitor (LMPP) and common myeloid progenitor (CMP) are further down the hierarchy, with the CMP preceding Erythrocytes and Megakaryocytes via the Multipotent Erythroid Progenitor (MEP). The GMP (Granulocyte-macrophage progenitor) population is thought to be more heterogenous, giving rise to monocytes, neutrophils, eosinophils, and dendritic cells. The Common Lymphoid Progenitor (CLP) gives rise to B cells, T cells, and Natural Killer cells. (Figure based on material from Laurenti *et al.*)

Mature bone marrow megakaryocytes appear anchored to bone marrow endothelial cells (BMECs) and to project proplatelet extensions into the lumen of bone marrow sinusoids [34, 35, 36, 37]. It has been suggested that interactions with BMEC adhesion molecules, stromal cell-derived factor 1 (SDF-1) and fibroblast growth factor 4 (FGF-4) have been implicated in an axis thought to induce the expression of very late antigen (VLA-4) on MKs and VCAM-1 on BMECs [38, 39, 40]. Recent work investigating the role of Cdc42 and RhoA demonstrates that the crosstalk of these systems drives transendothelial proplatelet projection[41].

Interestingly the paradigm of MK and HSC distribution in the bone marrow has also recently been challenged. It has been previously thought that HSCs primarily reside within an osteoblastic niche wherein they expand[42]. Upon lineage commitment, these cells were thought to migrate to bone marrow sinusoids, where they subsequently mature as MKs and eventually produce platelets[42]. More recently however, an elegant whole bone lightsheet approach has shown an unbiased distribution of mature and immature MKs throughout the bone marrow[43]. This work describes a spatial regulation of MK distribution throughout the marrow on the basis of an extensively vascularised space which restricts migration and movement, implying that MKs mature in these tightly defined microenvironments are in constant proximity to blood vessels [43].

Interestingly, while MKs culture *in vitro* have been shown to develop proplatelet extensions (particularly in murine foetal liver cells), the efficiency and scale of this is markedly reduced when compared to the thousands of platelets produced by MKs *in vivo*. This remains a significant obstacle in the *in vitro* study of platelet production, as well as the burgeoning field of artificial platelet generation, where even the most sophisticated bioreactors fail to emulate the thousands of platelets produced by MKs physiologically[44, 45, 46, 47, 48, 49]. Work by Stegner *et al.* and others implies that there is a critical role for proximity to endothelial sinusoids in both MK maturation and platelet production.

Regardless of the mechanisms by which HSCs commit to the megakaryocytic lineage,

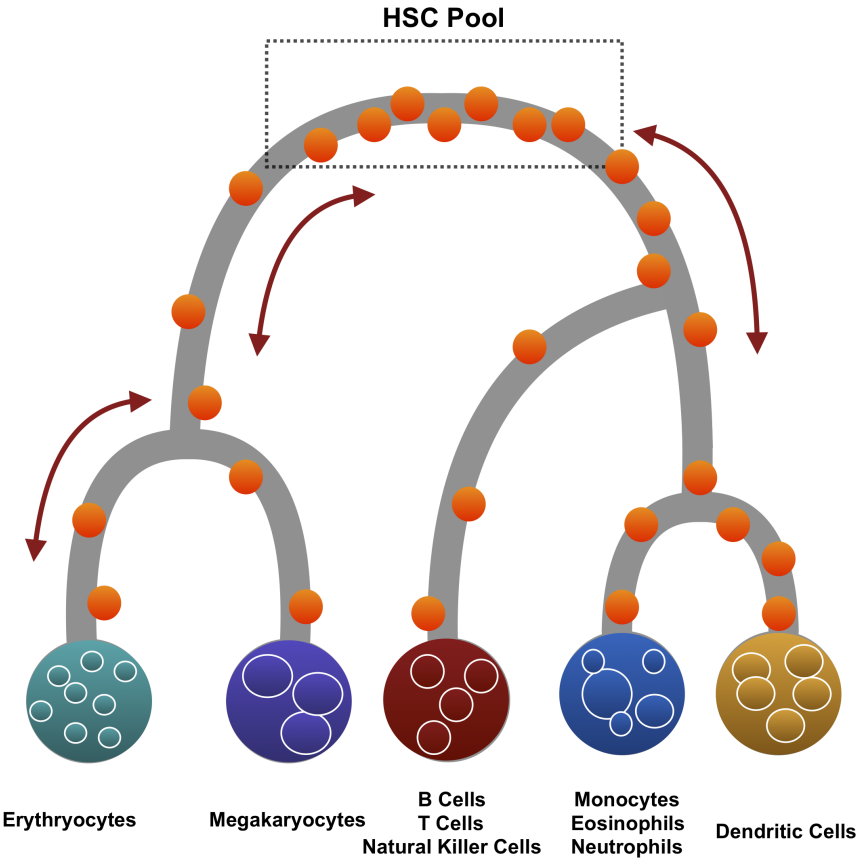

Figure 1.3: **Redefining haematopoeitic hierarchies as a continuum based on single cell transcriptional profiling.** The dogma of a restricted, lineage tree based hierarchy as defined by years of flow cytometry has been challenged by single cell transcriptomics and RNA seq. An emerging model suggests that HSCs exhibit a continuum of differentiation, suggesting a plastic system capable of responding to stress to rapidly produce haematopoeitic cells as needed. Figure based on material from Laurenti *et al.*)

the development and maturation of megakaryocytes is a remarkably specialised process involving the differential expression of critical genes, proteins, and transcription factors. The result is the development of large, polyploid cells able to produce and maintain the normal physiological levels of platelets essential for healthy haemostasis.

In the classic and most well defined hierarchy, a multipotent, committed colony forming progenitor precedes early megakaryocytes - this cell is the CFU-GEMM, or Colony Forming Unit Granulocyte Erythroid Macrophage Megakaryocyte capable, as the name suggests, of differentiating into any one of the three named lineages (erythroid, basophilic, and erythroid)[50, 51, 52]. On commitment to the MK lineage, the burst forming unit megakaryocyte forms (BFU-Meg)[52, 53]. These early progenitors express a mix of both HSC and MK factors, including CD34 for the former and CD41a for the latter (Figure 1.4).

Megakaryocytes within the bone marrow present as one of three distinct morphologies. The first, the promegakaryoblast, varies in size and is dominated by a nucleus up to five times the size of the surrounding cytoplasm[54]. The megakaryoblast (or stage I megakaryocyte) typically features a kidney shaped nucleus, and like its precursor the promegakaryoblast, also appears basophilic in stained bone marrow preparations[54]. The stage II megakaryocyte or promegakaryocyte is typically substantially larger than its precursors (between 20-80μm) and appears less basophilic with distinctive developing granules[54]. As megakaryocytes develop, they accumulate distinctive characteristics which are hallmarks of mature, platelet producing cells[55, 56]. These cells begin to lose markers of multipotent stem cells like CD34 and CD45, and accumulate lineage specific markers (CD41, CD42, CD61), with their differentiation driven by the presence of critical cytokines like TPO, IL-3, IL-6, and IL-11 (Figure 1.4).

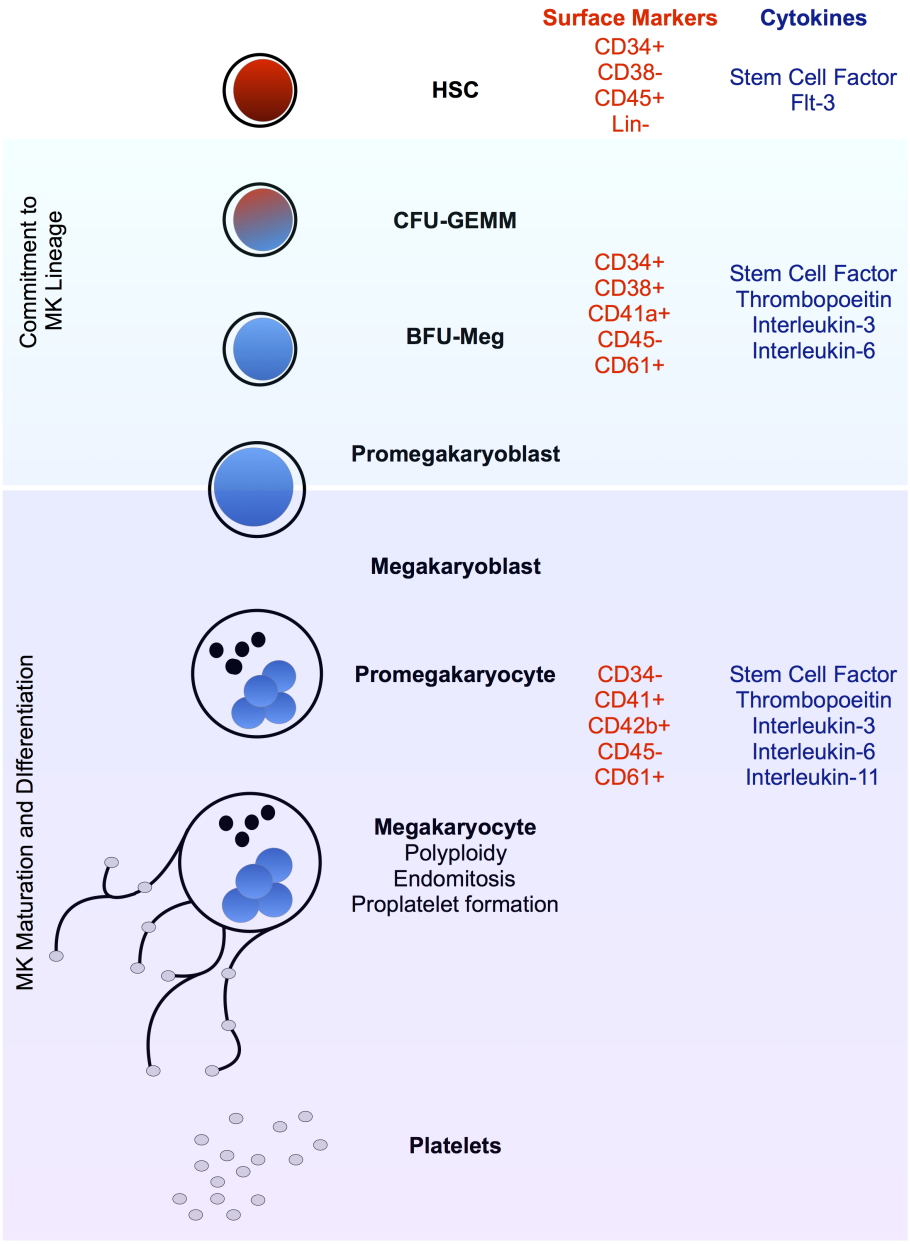

Figure 1.4: **Lineage commitment and maturation of megakaryocytes.** From a pool of self-renewing, multipotent haematopoeitic stem cells a multipotent colony forming unit (CFU-GEMM) develops before committing to the MK lineage through formation of the burst-forming unit megakaryocyte (BFU-Meg). These cells express both markers of stem cells and the MK lineage (CD34+, CD41a+), and their differentiation is driven by SCF, TPO, IL-3 and IL-6. As the colony forming unit develops into early megakaryoblasts, there is a loss of stem cell markers like CD34 and an accumulation of MK lineage specific surface proteins (e.g. CD61, CD42b), and differentiation is increasingly driven by TPO and IL-11.

1.2.2 Cytokines and the transcriptional regulation of megakaryopoeisis

As indicated in the previous section, megakaryocyte differentiation is heavily driven by the cytokine thrombopoeitin (TPO) and its binding to MPL, its receptor[57, 58, 59, 60]. Upon binding its ligand, MPL dimerises and induces the autophosphorylation of JAK2 (janus kinase 2), which subsequently phosphorylates an array of substrates to activate signalling pathways including PI3K (phosphoinositol-3 kinase) and MAPK (mitogen-activated protein kinases (Figure 1.5)[57, 58, 59, 60]. Over the last two decades, a host of downstream effects and effectors have been discovered, with more mechanistic consequences of TPO activation being reported[61].

Ultimately the effect of TPO is the induction of expression mechanisms which drive the differentiation and maturation of megakaryocytes. Interestingly, while it is the master regulator of megakarypoeisis and the subsequent maturation and development of megakaryocytes, TPO is not critical to platelet production - in fact, some evidence suggests that the presence of TPO may inhibit proplatelet production *in vitro* [62].

The cloning and purification of TPO led to the development of *in vitro* methods of MK culture and proliferation - significantly advancing the field[63, 64, 65]. Megakaryocytes derived from human CD34+ cells were shown to mature and form proplatelet extensions when cultured with TPO[63]. Interestingly work on TPO-deficient mouse models led to the discovery of a role for this cytokine in haematopoeitic stem cells. These mice not only demonstrated the thromboctyopenia that one would expect, but also a reduction in cell number across all haematopoeitic lineages[66, 67, 68]. This began work which would go on to establish TPO as a critical regulator of the haematopoeitic niche, and absolutely vital for the development and proliferation of megakaryocyte progenitors[61].

Other cytokines play a role in megakaryopoeisis - and these include SCF (Stem Cell Factor), Flt-3, Interleukin-6 (IL-6), and Interleukin-11 (IL-9). Optimising these cytokines

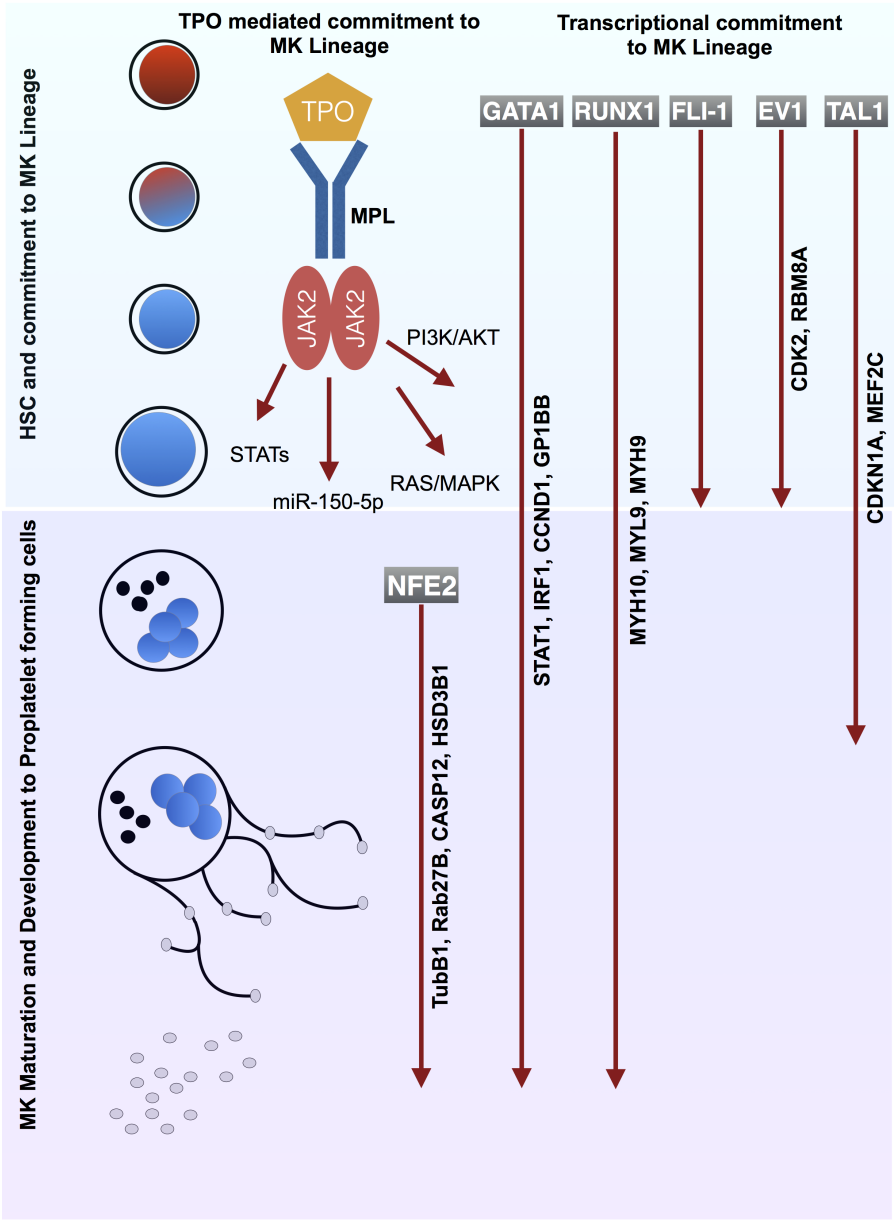

Figure 1.5: **Lineage commitment and maturation of megakaryocytes through TPO signalling and the actions of key transcription factors.** Thrombopoeitin signalling through its receptor cMPL is critical to the differentiation and maturation of megakarycoytes. TPO signalling initiates a cascade of downstream signalling events through the autophosphorylation of JAK2. These include activation of STAT1, STAT5, PI3K/AKT, and RAS/MAPK. The expression of key transcription factors is also an essential regulator of gene function in megakaryopoeisis. GATA1 and RUNX1 are expressed throughout MK development and maturation, while FLI-1, EV1-1, and TAL-1 are expressed during earlier stages. NF-E2 is expressed during the terminal maturation steps in which megakaryocytes begin to form proplatelet extensions, driving the expression of key markers of mature MKs e.g. *TubB1*, Rab27B, CASP12.

has been shown to affect the purity of primary MK cultures and the fraction of mature cells. Recent work by Machlus *et al.* identified a role of the cytokine chemokine ligand 5 (CCL5) which, upon binding to its receptor CCR5 triggers signalling through phosphorylation of the protein BAD. The resulting pro-survival signal may drive increased endomitosis and ploidy[69].

Differential gene expression and the resulting changes in protein expression are a key aspect of differentiation and the development of specialised cell types with specific morphologies. A number of mechanisms allow for the fine control of which genes are expressed and when - amongst them is the use of transcription factors (TFs), proteins which bind to specific DNA sequences to control the rate of transcription and thereby regulate gene expression.

A number of transcription factors have been shown to be critical to haematopoeisis and megakaryopoeisis - and indeed, the manipulation of these proteins has been used to forward programme stem cells to generate large numbers of megakaryocyte like cells[70]. Of these, the GATA family, RUNX1, Tal-1, EVI-1, and NF-E2 are known to have an important role in MK development and maturation(Figure 1.5)[71]. Perhaps unsurprisingly, deficiencies in a number of these TFs have been associated with inherited thrombocytopenias, and these will be discussed in section 1.4.

The zinc-finger transcription factors GATA1 and GATA2 are members of the GATA family of TFs, characterised by their binding of the T/A (GATA) A/G consensus sequence[71]. GATA2 has been established as key to the maintenance of HSCs, however it is still expressed in MKs, and therefore thought to have some overlapping function[72]. GATA1 is expressed in MKs (as well as erythroid and mast cells), and the loss of this TF reduces MK maturation by limiting cytoplasmic maturation and polyploidization. A variant of a form of inherited thromboctyopenia, Bernard-Soulier syndrome, features a mutation in the GATA-binding motif of the *GP1BB* promoter (Discussed further in section 1.4.2).

RUNX1 belongs to the RUNT family of TFs, and is considered a master regulator of haematopoeisis[71]. The loss of RUNX1 in MK lineages results in a phenotype similar to

the one observed where GATA1 is targeted - namely a reduction in polyploidization and maturation. RUNX1 defects are associated with thrombocytopenias and a predisposition to malignany (acute myeloid leukaemia)[73, 74]. Patients with RUNX1 mutations and their associated familial platelet disorders present with abnormal expressions of MPL and myosin light chains 9 and 10 (MYL9, 10)[75].

FLI-1 (Friend leukemia virus integration-1) works alongside GABPA (GA binding protein transcription factor alpha subunit), the latter binds genes associated with early MK development (GPIIB and MPL), while the former binds to both early and mature genes (GPIBA, GPIX, and PF4)[71]. The ratio of FLI-1 to GABPA increases with maturation driving the expression of genes associated with mature MKs. In patients with an inherited deletion of the 11Q chromosome and a subsequent hemizygous loss of FLI-1, MK differentiation is arrested with abnormally small cells (microMKs) observed in bone marrow smears as well as a thrombocytopenia and bleeding symptoms[76].

Depletion of T-cell acute lymphocytic leukemia 1 (TAL-1) in haematopoeitic cells results in a decrease of erythroid and megakaryocytic cells[77]. Cyclin-dependent kinase inhibior 1A (CDKN1A) is a key target of this transcription factor, and is aberrantly over-expressed upon a decrease of TAL-1 resulting in a reduction in the proliferation, ploidy, and maturation of MKs[78]. There is an equivalent reduction in platelet number, but no decrease in platelet function in this model. To confirm the role of TAL-1 as a regulator of CDKN1A, a knock-down of this target restores endomitosis in MKs[71, 79].

Nuclear factor Erythroid 2 (NF-E2) is a vital regulator of erythroid and megakaryocytic development (with roles in the myeloid lineage as well)[80]. Knock-outs of NF-E2 result in a complete absence of platelets despite an increase in the number of megakaryocytes[81]. This work highlights the role of NF-E2 as a terminal regulator of platelet production. Direct targets of this transcription factor include the *TubB1*, *Rab27b*, *CASP12*, and *HSD3B1* genes[82, 83, 84]. Each of these targets has an established role in platelet production, with *TubB1* for example

encoding the β-tubulin isoform specific to platelets and megakaryocytes (discussed at length in this thesis).

While each of these transcription factors are critical in their own right, and so have links to inherited thrombocytopenia and platelet disorders, they are part of complex regulatory networks in which they act to regulate one another. These are detailed by Tjissen *et al.* in their excellent review on the role of transcription factors in megakaryo- and thrombopoeisis[71].

1.2.3 Endomitosis and polyploidization

Mature megakaryocytes undergo a unique process by which they accumulate their hallmark polyploidy referred to as endomitosis [85, 86, 87]. This is a process of repeated, arrested cell division which results in the unique, lobed nuclei typical of mature megakaryocytes[85, 86, 87]. The extent of polyploidization varies extensively, particulary *in vitro*, ranging in multiples of a diploid chromosome content (4N, 8N, 16N, 32N, 64N) as a result of two to six endomitotic cycles. Polyploidization is likely a means by which to achieve gene amplification and a subsequent increase in mRNA and protein production in preparation for thrombopoeisis[88].

An area of significant interest is the method of deregulation which allows megakaryocytes to undergo an abnormal amplification of DNA and aborted mitosis, processes which are tightly regulated in other cell types to avoid malignancy. It has been suggested that a reduction of mitosis promoting factor (MPF) is involved in endomitosis[89]. In a typical cell, the synthesis of cyclin B and its subsequent binding to Cdc2 results in an active MPF complex whose kinase activity is required for mitosis. However, while cell line studies have implicated inactive Cdc2 or the absence of cyclin B, work in primary megakaryocyte cultures has shown normal levels of both - and thus active MPF[90, 91, 92].

1.2.4 Invaginated Membrane System (IMS)

Following polyploidization, megakaryocytes also demonstrate a unique processs of cytoplasmic maturation which is critical to their role as platelet progenitors. This results in the accumulation of material which will ultimately be passed on to platelets - including organelles, platelet surface receptors and other proteins. The development of the invaginated membrane system (IMS - formerly referred to as the demarcation membrane system or DMS) results in a characteristic membranous network which was originally thought to form platelet territories[93]. More recently, the IMS has been suggested to act as a source of membrane during the process of proplatelet extension[94].

Recent work by Eckley *et al.* has shown that the plasma membrane folds at specific points, to generate a peri-nuclear early IMS (pre-IMS) which expands through the addition of material from the golgi and endoplasmic reticulum[95]. The IMS is exposed to the extracellular environment allowing for labelling with a variety of tracers[96, 97].

1.2.5 Platelet granules

Secretory granules (α and dense granules) are critical to the vital roles platelets play in haemostasis and thrombosis. As such, granule generation and their subsequent trafficking and packaging into platelets is a critical component of megakaryocyte biology and platelet production. α-granules are typically between 200 and 500 nm in size, and these contain proteins critical to platelet adhesion and activation in the event of vascular injury[98]. These golgi derived secretory vesicles are typically evident in early stage megakaryocytes, and accumulate their contents both from endogenous production of protein and uptake. Platelet factor-4 (PF4) and von Willebrand Factor (vWF) are examples of endogenously generated α-granules proteins, while fibrinogen is a classic example of an endocytosed protein whose uptake is mediated by αIIbβ3[99, 100, 101]. α-granules also contain a number of platelet surface receptors including αIIbβ3, P-Selectin (CD62P), and CD36. As multivesicular bodies typically

contain P-selectin, vWF and other components typical of α-granules, they are thought to be precursors to these organelles.

Dense granules are typically 250nm in size and are so named due to their typically electron-dense cores as evidenced by electron microscopy. These are packaged with small molecules like serotonin, calcium, ATP, ADP and catecholamines[98]. Italiano *et al.* proposed a model of differential granule release based on the observation of non-overlapping granule proteins (VEGF as a pro-angiogenic factor localised to specific granules)[102]. Interestingly this model was contested by Kamykowski *et al.* in 2011 in a comprehensive super-resolution study of granule content in which no distinctive, non-overlapping granules were reported[103]. However the authors do report sub-clustering of specific proteins within granules, which raises the possibility of differential release.

Granules are known to be intimately associated with microtubules, as such microtubule associated motor proteins are likely to be critical for the transport of these organelles down the length of the proplatelet shaft. This is detailed further in Chapter 4 which discusses the role of microtubules in platelet production and megakarypoeisis in more detail.

Emerging techniques for rapid, volumetric live cell imaging techniques promise to shed new light on the biology of granule production, transport, and release. The hope is that these methods will supplement previous work in the field which has primarily applied electron microscopy.

1.3 Platelet production

Megakaryocytes have ultimately one terminal function, the generation of platelets in the numbers needed to maintain haemostasis and effectively respond to vascular injury. The exact mechanism by which megakaryocytes produce platelets *in vivo* remains controversial. This is largely due to the difficulties presented by the bone marrow niche in which megakaryocytes mature and form platelets, and the differences in the behaviour of these cells when removed

from this environment and studied *in vitro*. Classic models of platelet formation and production include cytoplasmic fragmentation, budding, and proplatelet formation.

The discovery of TPO and its MK specific receptor c-Mpl has been instrumental in driving MK and platelet research, allowing for the *in vitro* culture of primary MKs. The mechanism of MK differentiation, and the subsequent maturation of these cells into platelet producing, polyploid cells has been intensively studied as a result. *in vitro* models include murine megakaryocytes flushed from bone marrow, murine foetal liver MKs, human cord and peripheral blood derived cells.

Early work looking at electron micrographs of MKs noted surface blebbing of a size comparable to platelets - leading to the suggestion that platelets are shed from the MK cell periphery[104, 105]. This suggestion was quickly undermined by thinner sectioning approaches which revealed that these structures lacked platelet organelles. Early observations of the extensive internal membrane system (see section 1.2.4) characteristic of mature MKs led to an alternative model which described a 'demarcation membrane system' thought to form defined platelet territories[106]. Platelet production was thought to occur as a result of cytoplasmic fragmentation along these territories[107]. While these regions are positive for platelet organelles and proteins, they lack a number of the distinct, defining morphological characteristics of the archetypal platelet - for example, the microtubule marginal band[108, 109].

The proplatelet model of platelet biogenesis remains one of the more convincing models due to a large number of imaging studies which report the generation of long, beaded projections both *in vitro* and *in vivo*[110, 43, 42]. This theory suggests that platelets are formed through the generation of long protrusions into bone marrow sinusoids which, in the presence of shear, bud off to form preplatelets and platelets. In this model, the IMS acts as a reservoir of membrane which is extruded during the formation of proplatelet extensions[111].

Imaging technologies have generated a large body of evidence supporting the proplatelet model - most significantly perhaps is intravital imaging work which has shown (albeit

at relatively low resolutions) the generation of long, thin protrusions through endothelial junctions[43, 42]. *in vitro* proplatelet extensions have been observed in mature MKs derived from a number of species - including human cells. In thrombocytopenic models a loss of proplatelet extensions has been reported as causative of low platelet counts.

In its current iteration, the proplatelet theory as studied by live and fixed cell imaging approaches describes a series of changes which occur over the course of several hours (between 4 and 10)[112]. This is shown in figure 1.6, a time lapse of murine foetal liver MKs which sponteneously form proplatet extensions *in vitro* upon maturation. The process begins with the polarisation of the cell: a movement of the polyploid, lobed nucleus to one end of the cell while the IMS and related structures are concentrated to the other[28]. Pseudopodia develop and elongate to form long, thin extensions with consistent diameters ranging between 2-4 microns. As evidenced by live cell imaging, these extensions undergo a remarkable and dynamic series of branching and bending events punctuated by beaded structures which eventually resolve as platelets. Ultimately, this process leaves the megakaryocyte as little more than residual nuclear material surrounded by a complex network of beaded proplatelet extensions[28]. A rapid retraction followed by the rupture of the extensions between platelet sized beads is thought to result in the release of platelet like particles. Interestingly these final stages remain difficult to visualise using classical approaches as the temporal resolution required over the full observation period is phototoxic in conventional light microscopy.

The dynamic and often dramatic rearrangement of the MK to form proplatelet extensions is, perhaps unsurprisingly, driven by unique cytoskeletal processes. Both microtubules (MTs) and actin play a critical role in the significant morphological changes needed to generate platelets. As this work will primarily focus on microtubules in platelet production, this particular cytoskeletal protein is discussed in more detail in the following sections - with actin and its associated proteins touched on briefly.

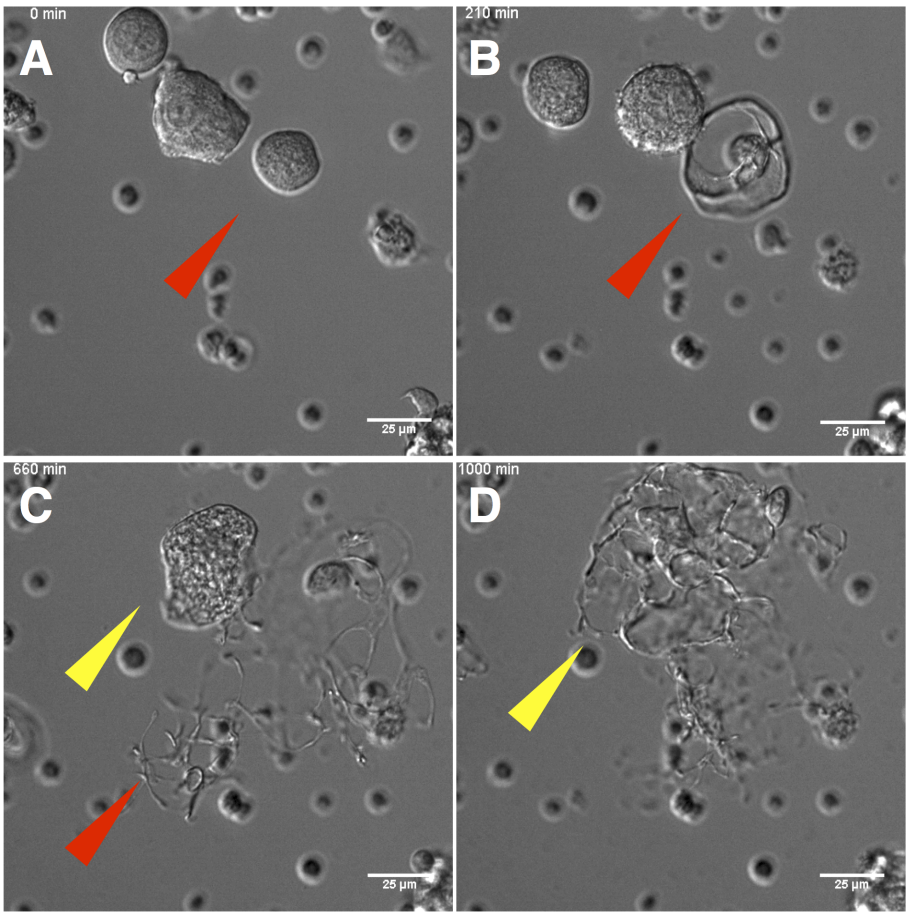

Figure 1.6: **Proplatelet formation in murine foetal liver megakaryocytes**. Foetal liver megakaryocytes rapidly produce proplatelets *in vitro* after 4 days in culture. (A) Large, polyploid cells dominate the culture at the beginning of the acquisition. (B) After 210 minutes a cell begins to expand, extensively reorganising its cytoskeleton and altering its morphology (particularly notable when compared to surrounding cells) . (C) At 660 minutes a series of long protrusions are evident, with distinctive beads along the length of each extension. Another cell in this field of view begins to re-organise. (D) At 1000 minutes the second cell has formed a network of long extensions. Images acquired live on a Zeiss Observer 7 system.

1.3.1 Microtubules in platelet production

Microtubules are a ubiquitous cytoskeletal protein with extensive roles in cell processes including cell division and vesicle transport. They are comprised of α and β tubulin heterodimers which assemble into protofilaments, these are further bundled into microtubules - hollow filaments approximately 25nm in diameter which exhibit uniquely contrary properties of dynamic stability[113, 114]. Microtubules are polarised, with bundles of MTs often exhibiting mixed polarities. The plus (+) end is defined by its exposed β-tubulin subunit, while the opposite (-) end is typified by an exposed α-tubulin. While MTs can polymerise in both directions, polymerisation is typically more rapid at the plus end.

Highly specialised cell types exhibiting unique, dynamic processes often achieve these functions through the regulation of microtubule function in what is referred to as the 'tubulin code'[113, 114]. This tubulin code is executed through the differential expression of cell specific tubulin isoforms, and the subsequent recruitment of MIPS (microtubule interacting proteins) and MAPS (microtubule associated proteins)to unique residues at key sites for post-translational modification[113, 114].

β-1 tubulin is a megakaryocyte and platelet specific isoform of tubulin encoded by the *TubB1* gene[115, 116, 84]. Expression of this β tubulin variant is specific to mature MKs, and as such defects in these protein have been linked to defective proplatelet formation and platelet function[84, 43]. The specificity and significant divergence of β-1 (20% at the amino acid level compared to other β-tubulin isoforms) has implicated this cytoskeletal protein as a critical part of the complex cytoskeletal machinery driving megakaryocyte and platelet function. Despite this, our understanding of how *TubB1* achieves its unique functions in MKs and platelets remains relatively poor.

TubB1 knock-out mice have shown a substantial reduction in platelet count (approx. 50 percent less in homozygous knock-outs), an increase in mean bleeding time (260s vs. 50s in wild type), reduced proplatelet formation *in vitro*, and a loss of the classic discoid shape

of platelets observed in wild type mice (instead becoming spherical)[82, 84]. The authors also observe poorer platelet activation in knock-outs when compared to wild type as evidenced by P-selectin labelling in response to thrombin. The detailed role of β-1 tubulin in human thrombocytopenias is discussed in Chapter 4.

Over the course of MK maturation microtubules evidence a substantial change in morphology and organisation. This is likely to correspond with the expression of both β1-tubulin and a number of other MK specific factors[117]. In earlier stage MKs microtubule assembly at centrosomes will stop and MTs move to the periphery of the cell where they form thick bundles in pseudopodia. As these thick structures begin to extend, MTs rearrange into linear arrays which are thickest at the portion of the extension nearest to the cell body (and become progressively thinner along the length of the proplatelet protrusion)[117].

Microtubule polymerisation continues throughout the proplatelet structure, however the rate of polymerization exceeds the rate of extension growth (10.2 μm/min vs a proplatelet elongation rate of 0.85 μm/min). This coupled with the observation that microtubules polymerise both towards the end of the proplatelet shaft and the cell body implies that polymerisation and proplatelet extension do not necessarily correlate. This is further supported by work by Bender *et al.* [112, 118]. The authors reported a dynein dependent mechanism of microtubule sliding analogous to systems observed in flagellated cells and axons [118, 119, 120, 121]. Dynein is one of the two motor proteins associated with MTs (the other being kinesin). Dynein is a (-) end directed motor protein which has also been reported as a critical player in the maintenance of the platelet marginal band at rest, and its expansion and coiling on activation[120, 121].

Bender *et al.* report that while inhibition of microtubule polymerisation does not impair proplatelet production (consistent with previous reports), the inhibition of cytoplasmic dynein (EHNA and Ciliobrevin A) significantly impairs proplatelet extension.

Interestingly, recent work has shown that artificial expression of β-tubulin VI in CHO

cells under the control of a tetracycline inducible promoter has a significant impact on microtubule assembly and function[115]. The authors show that an abnormal bundling of microtubules is observed at the cell periphery mimicking the formation of the marginal band, and that this is linked to expression level in the stable cell lines generated. Similarly, clones expressing β-tubulin VI demonstrated aberrant spindle assembly and cell cycle progression (prometaphase arrest), with an increase in microtubule acetylation. This work implies that β-tubulin expression is key to driving endomitosis and the unique microtubule localisation exhibited by megakaryocytes during and after endomitosis.

Platelets require an abundance of specialised proteins packaged in granules to execute their haemostatic functions - therefore the process of platelet production involves the trafficking of organelles, granules and their contents along proplatelet extensions and into what will be a terminal, mature cell[122]. These components have been shown to be in direct contact with microtubules and are driven in both directions along proplatelet shafts through the action MT associated motor proteins[28].

Kinesin (a (+) end directed MT associated motor) appears to be related to the distribution of organelles and granules - and is thus likely to be the motor responsible for transporting nascent granules along the length of the proplatelet shaft. Recent work by Adam *et al.* implicates Kinesin-1 as the motor involved in platelet secretion[123]. In their study a conditional Kif5b knock-out was used to assay the effects of the loss of the major kinesin-1 heavy chain in the MK lineage - the authors report no change in the morphology of platelets, their number, or the distribution and number of granules[123]. They do however report a reduction in platelet aggregation *in vitro* inconsistent with the lack of phenotype otherwise observed.

The role of MTs and their motors in the context of platelet activation and the broader literature are discussed in more detail in the introduction to Chapter 4.

1.3.2 Actin in platelet production

Globular actin is assembled into filamentous structures (microfilaments) which are critical in cell motility and division amongst a host of other functions. In megakaryocytes, actin branching drives proplatelet bending and facilities the generation of force through interactions with its motor, non-muscle Myosin. MKs treated with actin inhibitors (e.g. latrunculin, cytochalasin) form long proplatelet extensions which lack their characteristic branched appearance[112]. Mutations in the heavy chain-of non-muscle myosin (MYH9) results in May-Hegglin anomaly, a condition which presents with a macrothrombocytopenia[124, 125].

Actin is regulated by a host of interacting proteins which help this ubiquitous cytoskeletal protein to achieve cell specific functions. In platelets and megakaryocytes these include WASp (Wiskott-Aldrich Syndrome protein), Profilin1A, Arp 2/3, and Cofilin. Wiscott-Aldrich syndome proteins are autoinhibited and activated by Cdc42 and PIP2 to allow for the binding and activation of the Arp2/3 complex (which nucleates actin). Human patients and WASp knock-out mice both demonstrate low platelet counts as a result of increased clearance with evidence of ectopic platelet release - although platelet function in these instances has been reported as normal[126, 127].

Profilin is an ADP binding protein which interacts with the ends of actin filaments. Mouse models deficient in profilin 1 exhibit a significant microthrombocytopenia[128, 129]. The Arp2/3 complex drives actin nucleation, with the ARPC1B subunit highly expressed in cells with a haematopoeitic lineage. A recent study by Kahr *et al.* reported that the loss of the ARPC1B isoform caused a complex condition featuring inflammatory disease, vasculitis, thrombocytopenia, and platelet dysfunction. In ArpC2 knock out mice a microthrombocytopenia is observed typified by an ectopic release of platelets within the bone marrow[130].

1.3.3 Spectrin in platelet production

While the majority of current and previous research into the cytoskeletal mechanisms driving platelet production focus on actin, microtubules, and their myriad interacting and regulating proteins, other components are likely to play a key role in the dynamic and extensive shape changes evident in platelet production[131, 132]. The membrane skeleton has been recently reported as important for elements of MK development. Proplatelets have been reported to feature a spectrin membrane skeleton, with non-erythroid subunits alpha-II and beta-II spectrin highly expressed in MKs (erythroid isoforms are expressed albeit at lower levels)[131, 132].

Treatment of MKs with a tetramer-disrupting peptide inhibits the formation of the IMS or proplatelets. When the same peptide is applied to detergent permeabilized MKs, proplatelet morphology is disturbed[131, 132]. This remains a nascent field, however the evidence to date suggests an important role for the spectrin tetramer in platelet production.

1.3.4 Platelet production and apoptosis

In many ways MK maturation and the process of platelet production resemble a directed form of cell death. Both pro- and anti-apoptotic factors have been described in MKs. Bcl-2 and Bcl-XL are anti-apoptotic factors found in early MKs, and when these factors are overexpressed in these cells the result is an inhibition of proplatelet formation[133, 134]. Work by Debili *et al.* investigating apoptotic caspases in cord blood derived MKs report a loss of proplatelet formation when these factors are inhibited [135]. In contrast however, knock-out mice deficient in Caspase-8, Caspase-9, and the pro-apoptotic factors BAK, and BAX report no reduction in platelet counts.

It is suggested that MK maturation is a product of the inhibition of the intrinsic pathway of apoptosis - with key survival proteins involved in the repression of intrinsic apoptosis. MCL-1 knock-out mice demonstrate no defects in platelet production, while treatment with

an inhibitor of Bcl-XL and Bcl-2 result in MK cell death. This evidence suggests that repression of apoptosis is required for platelet production[133, 134].

1.3.5 Platelet release

The production of platelets ultimately terminates in the release of platelets into the blood-stream - the observation that the anucleate fragments launched into the vasculature are typically larger than mature platelets implies that an intermediate stage precedes the final product[136, 137]. These 'preplatelets' are thought to precede the final stages of platelet formation which occur in circulation. 'Preplatelets' are reported as discoid cells up to 10μm in diameter which form a characteristic barbell-shape prior to fission into platelets. This final step is likely driven by MT associated motors - and indeed the macrothrombocytopenias discussed in Chapter 4 are thought to occur as a result of a loss of this key MT mediated behaviour[136, 137].

Intravital microscopy has provided a remarkable insight into the release of platelets *in vivo*, with a number of excellent works documenting the generation of proplatelet extensions into the lumen of bone marrow sinusoids[42, 43]. Recent advances in imaging technologies are likely to improve on the resolution of previous works and provide further insights into this process.

1.4 The molecular basis of inherited thrombocytopenias (ITs)

Inherited thrombocytopenias (IT) are a class of hereditary diseases characterized by abnormally low platelet counts (below the normal range of 150-450 x10^9/L), often manifesting as bleeding diathesis [138, 139, 140]. With the advent of improved, high throughput sequencing technologies, ITs have been revealed to be complex inherited disorders resulting from a diverse range of novel mutations (Figure 1.7). Moreover, these often unique variants have, upon further investigation, yielded novel insights into platelet and megakaryocyte (MK) biology on a molecular level [141, 139, 140].

ITs are often difficult to diagnose, firstly as rare diseases which, due to variation in their molecular causes, can present with a range of symptoms[141]. The classic hallmark of ITs, an abnormally depleted platelet count, is one which is often only diagnosed after a patient has been sufficiently challenged [141]. Interestingly Morowski *et al.* show that mice can tolerate significant decreases in platelet count before presenting with a bleeding phenotype[142]. The authors systematically deplete platelet counts using antibodies and subject treated mice to a number of challenges, showing a surprising tolerance of reduced platelet counts in terms of functional assays [142]. This work highlights the need for human, *in vitro* based assays when interrogating human mutations.

Similarly unique patient mutations can have markedly different effects on protein function, ranging from a complete loss to truncations with functional consequences, with possible species dependent variations. These can cause thrombocytopenias due to defective platelet production, clearance, or MK maturation. Unpicking the mechanistic effects of mutations can be quite complex, as in many instances there are further characteristics of particular thrombocytopenias which compound the initial mutation. For example, as detailed later in this chapter, particular thrombocytopenias are associated with enlarged platelets or defects

in granule content and release.

The last decade has seen the discovery of an unprecedented number of novel genes implicated in ITs due to the emergence of Next Generation Sequencing (NGS) technologies, most notably the Whole Exome Sequencing (WES) approach [139]. To date, 31 genes responsible for 27 distinct inherited thrombocytopenias have been reported [141, 139, 140]. These molecular approaches have linked mechanistic causes with patient phenotypes, prominent examples of which include myosin heavy chain 9 (MYH9)-related and Bernard-Soulier disease[143, 144, 145, 146].

Future challenges are in identifying novel genes and patients whose diathesis are a result of different variants working in combination to affect megakaryopoeisis and platelet production[145]. Typically, defects implicated in ITs can be grouped according to their effects on the differentiation of platelet progenitors, MKs, the maturation of MKs, and platelet formation, or in a minority of cases, in genes and proteins not directly involved in the process of generating platelets[139, 140].

This section will briefly discuss a range of inherited thrombocytopenias and their respective molecular causes. *TubB1* and previous reports of thrombocytopenias which arise as a consequence of its mutation are excluded from this section for a more comprehensive discussion in Chapter 4.

1.4.1 Grey Platelet Syndrome

The loss of α-granules results in the classic gray appearance of platelets on blood smears of patients affected by Grey Platelet Syndrome (GPS) [147, 148]. Patients diagnosed with GPS will typically present with macrothombocytopenia, bleeding diathesis, splenomegaly, and bleeding diathesis[149].

Both platelets and megakaryocytes in GPS present with a loss of α-granules, while dense granules and lysosomal compartments appear unaffected. As secretory vesicles carrying an abundance of factors promoting adhesion and wound healing upon platelet activation,

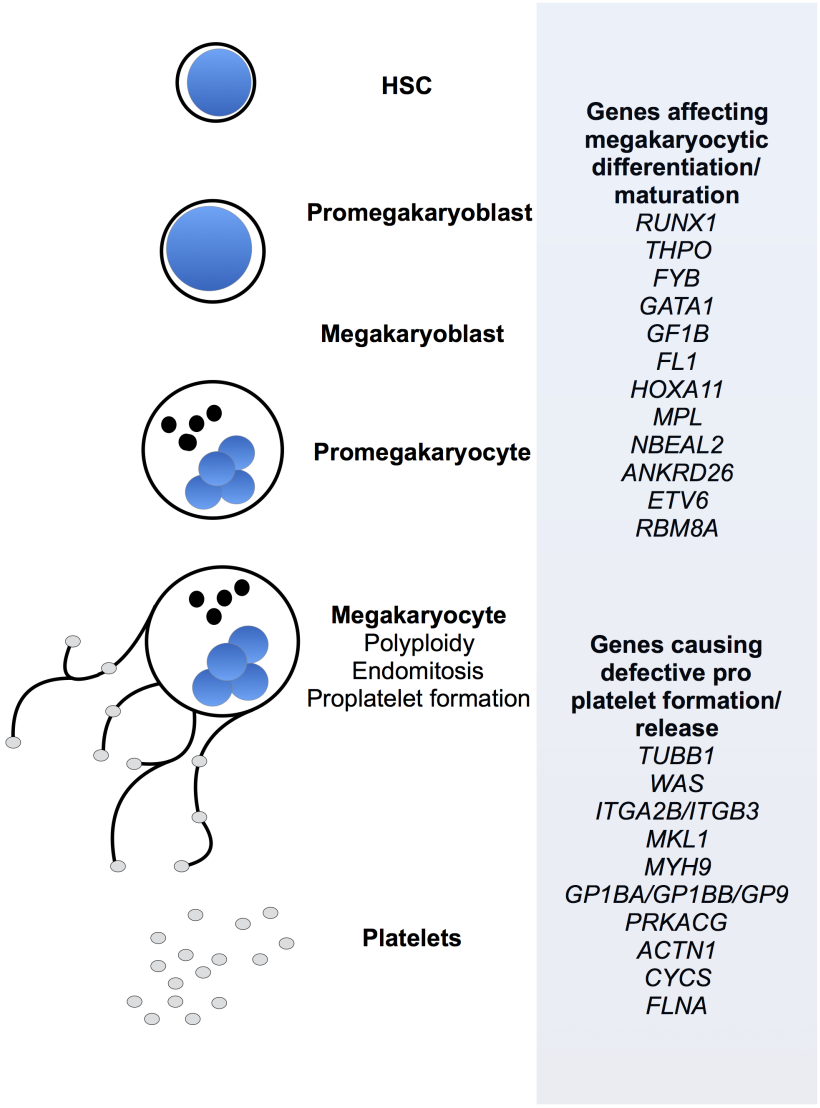

Figure 1.7: **Megakaryocyte differentiation is a stepwise process over which MKs accrue hallmark features.** Arising from the haematopoeitic stem cell (HSC), the promegakaryoblast is a larger cell with a low nucleus to cytoplasm ratio. As the cell develops becoming a megakaryoblast, it forms a distinctive kidney shaped nucleus. Further differentiation results in the formation of the promegakaryocyte, a cell type which begins to accumulate granules and the loss of the previously basophilic appearance of the cell. A mature megakaryocyte is the final result of this differentiation process. This cell type demonstrates the characteristic polyploidy which occurs as a result of endomitosis and can go on to form platelets. This maturation process is defined by the expression of a series of transcription factors implicated in thrombocytopenias. Genes associated with inherited thrombocytopenias can be broadly split into two key groups - genes affecting MK development, differentiation, and maturation, and those resulting in defective proplatelet formation or release. (Figure adapted from Johnson *et al.* [140].)

GPS patients will as a result present with bleeding[149]. Interestingly proteins synthesised by MKs are particularly deficient in these cases (e.g. PF4, β-thromboglobulin), while other components of these vesicles are less reduced.

GPS is caused by variations in the gene *Nbeal2* (Neurobeachin-like 2) which encodes a protein of the same name which is found in the platelet endoplasmic reticulum[150, 151]. *Nbeal2* null mice demonstrate the same macrothrombocytopenia with a loss of α-granules, as well as a delayed maturation of MKs and a reduction in their ploidy[152].

1.4.2 Bernard-Soulier Syndrome

Bernard-Soulier Syndrome (BSS) is a rare inherited bleeding condition which presents with defects of the GP1b-IX-V complex[153]. Typical symptoms include reduced platelet numbers, giant platelets, and bleeding tendency[153]. Loss of function in this receptor results in poor binding with von Willebrand Factor, an event critical for a healthy haemostatic response[154].

The GP1b-IV-V complex is formed from the expression of four genes (*GP1BA, GP1BB, GP9, and GP5*) as MKs mature in the bone marrow. Mutations causing BSS are spread over these four genes, with the exception of GP5 (in which no mutations causing BSS have been reported thus far)[154]. Due to the range of possible causative mutations, BSS can present with varying degrees of severity and a host of symptoms[154].

1.4.3 MYH9-Related disease

MYH9 encodes Myosin 9, the heavy chain of a key non-muscle myosin IIA heavily involved in platelet production[139, 140]. Interestingly mutations in this gene have been shown to increase the production of proplatelet extensions, suggesting a role for this protein in the terminal phase of platelet fission - the loss of which causes thrombocytopenia[155, 156].

MYH9-Related diseases include the May-Hegglin anomaly, Sebastian, Epstein and Fechtner syndromes. Each condition shares key clinical features, namely the formation of a small

number of enlarged platelets (macrothrombocytopenia) and leukocyte inclusions known as Dohle bodies[157].

1.4.4 Paris Trousseau type thrombocytopenia

Paris-Trousseau thrombocytopenia is a bleeding diathesis demonstrated by individuals with a terminal deletion of chromosome 11, a rare condition referred to as Jacobsen syndrome[158]. In these patients, a reduced platelet count is accompanied by platelets with giant α-granules and abnormal responses to thrombin [159, 160, 161]. These individuals will also feature two distinct subsets of megakaryocytes, one normal and the other abnormally small [158, 159, 160, 161].

Amongst the genes lost by the Jacobson deletion is the transcription factor FLI-1, a key regulator of megakaryocyte development described in a previous section [158, 159, 160, 161]. The loss of FLI-1 has been modelled in mice and human stem cells and is likely to be causative of Paris-Trousseau thrombocytopenia in Jacobsen patients.

1.4.5 Wiskott-Aldrich Syndrome

Actin polymerisation is a key component of the morphological changes associated with both platelet production and function. The actin nucleation promoting factor WASp, encoded by the *WAS* gene and expressed in haematopoeitic cells, has a critical role in platelet production and function (discussed in more detail in section 1.3.2)[162]. Abnormal *WAS* expression is associated with Wiskott-Aldrich Syndrome, a rare X-linked condition which includes a low platelet count (X-linked thrombocytopenia) amongst other symptoms (e.g. micro platelets, eczema, immune deficiency)[162].

Interestingly the microthrombocytopenia observed in Wiskott-Aldrich syndrome is also evident in the recently reported ARPC1B disorder, supporting the role of actin nucleation (and subsequent branching) in the production of healthy platelets[163].

1.4.6 von Willebrand disease

von Willebrand Factor (vWF) is a vital mediator of platelet function in haemostasis and thrombosis, the loss of this protein results in a series of conditions known as von Willebrand Diseases (VWD) [164, 165, 166, 167]. vWF binds to the N-terminus of glycoprotein Ibα through its A1 domain when the protein is exposed in the subendothelium [164, 165, 166, 167].

In types 1 and 3 of VWD, a deficiency or complete lack of the protein cause the bleeding condition, however in type 2B a functional deficiency is responsible for the resulting diathesis[164, 165, 166, 167]. In VWD type 2B, mutations in the A1 domain of the protein result in a gain of function phenotype, wherein circulating vWF spontaneously binds to platelet GP1bα.

Interestingly a dysfunction in megakaryopoeisis has been reported in VWD type 2B by Nurden *et al.*. [168]. A family with a substitution mutation in the A1 domain of the protein show poorly differentiated megakaryocytes from peripheral blood cultures[168].

1.4.7 The UK Genotyping and Phenotyping of Platlets (GAPP) project and challenges in the study of inherited thrombocytopenias

With the importance of defining novel mutations responsible for ITs in mind, the Genotyping and Phenotyping of Platelets (GAPP) study was created with the intention of cataloguing and investigating platelet phenotypes and genotypes[169, 170]. With samples from patients with suspected thrombocytopenia and functional defects across the UK, whole exome sequencing (WES) has identified numerous novel defects in patients[139, 140]. A recent example is the identification of schlafen 14 (*SLFN14*) in patients with a history of thrombocytopenia, excessive bleeding, and platelet secretion defects[171].

12 patients from 3 unrelated families with similar phenotypes, namely a dominant

inheritance pattern, enlarged platelets, thrombocytopenia, and decreased ATP secretion, were found to harbour missense mutations within the GTP/ATP-binding region of *SLFN14*[171]. In this instance, not only has the underlying molecular cause of the patient phenotype been discovered, but the importance of this particular protein in platelet production has been highlighted. The identification of *SLFN14* is a prime example of how investigating patient mutations through the GAPP project can lead to an improved understanding of ITs and platelet biology, and current work is focussing on dissecting the role of *SLFN14* in platelet formation.

The work performed thus far on *SLFN14* also highlights a significant obstacle in the study of platelet disorders, namely the need for enough patient sample for the characterisation and study of novel defects. The advent of CRISPR-Cas gene editing has created a previously unprecedented opportunity for the generation of cells harbouring specific mutations of interest. Applying this gene editing tool to cell lines and stem cell derived MKs can overcome the historical limitations applied to such studies by the finite amounts of patient sample available.

1.4.8 Aims

The primary focus of this PhD has been the development of methodologies to apply these technologies to the study of novel variants and genes revealed by the GAPP project. In the absence of patient material and cell lines which mimic platelet production, there is a need for the development of reliable iPSC differentiation models to yield mature, proplatelet producing cells for phenotypic screening.

In the following chapters the use of CRISPR for the endogenous labelling of genes of interest is reported, and the development of a method of single molecule imaging through CRISPR is validated. Attempts to adapt these knock-in methods for CRISPR in iPSC are discussed, alongside the adaptation of protocols for the generation of mature, proplatelet forming MKs from induced pluripotent stem cells.

Finally, investigations into the role of post-translational modifications of *TubB1* are reported with a number of novel insights into how the expression of the MK and platelet specific isoform of β-tubulin drives the unique morphologies and functions of these cells.

Chapter 2

Development of CRISPR-PALM

In this chapter the development of a method of single molecule imaging through the endogenous expression of fluorescent proteins is reported. One of the aims of this project is to establish protocols for CRISPR mediated knock-in of donor DNA. Applying CRISPR-Cas9 editing to insert fluorescent tags offers an easy-to-assay means of developing these protocols. Extending this approach to insert photoswitchable tags offers a novel application of this gene editing approach to address a number of long standing issues in the field of single molecule localisation microscopy. Finally, the protocols developed through these experiments should translate to the introduction of patient SNPs through CRISPR as both single base pair and reporter knock-in rely on the same homology directed repair (HDR) pathway.

CRISPR mediated knock-in also offers a number of distinct advantages in the study of novel genes implicated in inherited thrombocytopenia. Where good antibodies are not available, endogenously expressed labelled protein offers a convenient and robust method of studying the distribution and function of genes of interest. Combined with stem cell approaches discussed in chapter 3, these methods offer a potential means by which candidate genes can be modelled and studied *in vitro*.

In this introductory section, super-resolution techniques are introduced with a particular emphasis on single molecule localisation microscopy (SMLM). The application of CRISPR-

Cas9 in the context of imaging techniques is discussed, and finally data pertaining to a novel method of single molecule imaging through knock-in, CRISPR-PALM (CRISPR-PhotoActivated Localisation Microscopy), is discussed.

2.1 Introduction: Sub-Diffraction or Super-Resolution Microscopy

Light microscopy has been a powerful tool in cell biology since its inception, and because of the continuous development of fluorescent reporters and immunofluorescence techniques, has become a mainstay of modern cell biology. Modern microscopes are virtually indistinguishable from their ancestors as ever advancing optical configurations, cameras, and illumination pathways allow for improved resolution, signal to noise ratio (SNR), and phototoxicity. Arguably the most significant advance in modern imaging technologies has been the development of super-resolution techniques - methods which overcome the classic resolution limit of light microscopes to provide nanoscopic insights into the workings of the cell.

Defined by Ernst Abbe as a function of the numerical aperture of an optical system and the wavelength of light, the diffraction limit has classically kept cellular imaging above a resolution limit of approximately 200-300 nm (Figure 2.1), restricting the detailed structural study of nanoscopic elements of the cell[172]. In light microscopy, resolution refers to the maximum distance at which two distinct objects or point spread functions can be resolved as individual components. In recent years, a number of approaches have been developed which circumvent the diffraction limit, allowing for sub-diffraction or super-resolution (SR) nanoscopy of living and fixed cells [172]. Super-resolved fluorescence microscopy techniques were awarded the Nobel prize in Chemistry in 2014 (the prize was awarded to William E. Moerner, Eric Betzig, and Stefan Hell, though these are by no means the only individuals responsible for the development of these techniques), and have since become widely spread

methods of interrogating biological samples on the scale of nanometers.

A number of approaches can be applied to 'break' the diffraction limit, each with its own benefits and drawbacks. In this introduction the most widely used super-resolution techniques are discussed with a particular emphasis on SMLM.

2.1.1 Structured Illumination Microscopy (SIM)

Structured Illumination Microscopy (SIM or SR-SIM) refers to imaging modalities which apply an interference pattern to samples during acquisition, this can be used to computationally extract high-frequency information from the sample which is typically discarded, and reconstruct images with a doubling of the lateral resolution limit and an axial improvement ranging between 150-300nm (depending on the implementation used)[173, 174, 175]. SIM offers relative simplicity compared to other SR techniques, no specialised reporters or sample preparation is needed, and so multiplexing (imaging multiple reporters) is straightforward when compared to methods like single molecule localisation microscopy (SMLM).

However, classical limitations of SIM have included slow acquisition speed and reconstruction artifacts[173]. In the early SIM systems reported by Gustaffson *et al.* a diffraction grating was iteratively rotated to achieve a modified excitation pattern - and so achieve the variations in phase needed for SR-SIM. This unfortunately, was a time consuming method as the grating would need be mechanically moved for each acquisition, often between as many as 9 to 15 times - ultimately limiting the applicability of the approach for live cell imaging[176]. Since this early implementation, an array of alternative methods have been developed to achieve the variations in angle and phase needed to generate SIM images - including spatial light modulators, digital mirrors, and electro- and acousto-optic devices. These approaches have led to progressively faster SIM, and as a result, live cell applications which benefit from the efficient use of photons in modern cameras[177].

As reconstruction is needed to extract the relevant information needed from SIM ac-

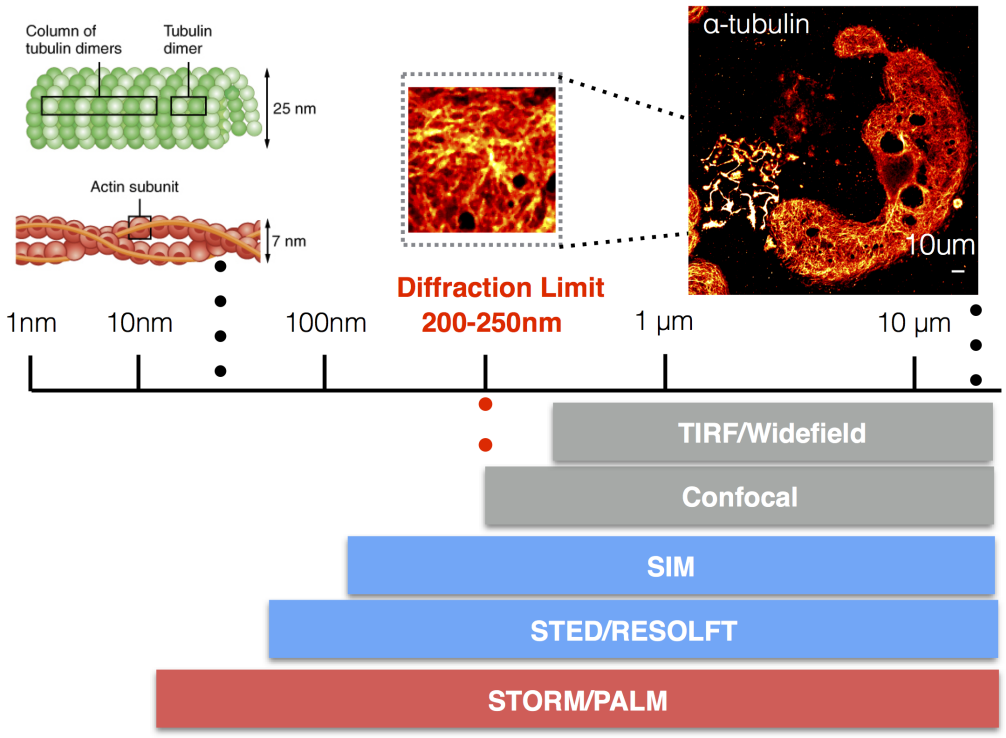

Figure 2.1: **The diffraction limit and super-resolution imaging modalities.** While whole cells are typically described on the scale of microns, most sub-cellular structures measure on a nanoscopic scale and receptors and other proteins are spatially organised in the tens of nanometers. 'Standard' imaging modalities like widefield or confocal microscopes are limited to the diffraction limit (200-250nm laterally), however an array of super-resolution techniques can provide insights on a scale of nanometers, which is much more appropriate for describing sub-cellular organisation and distribution. Cytoskeletal structures are an excellent example of cellular components which exist on a scale an order of magnitude beneath the classical diffraction limit. Microtubules and actin filaments have diameters of approximately 25nm and 9nm respectively, substantially lower than the resolution limit of conventional light microscopy. While techniques like SIM and STED can offer substantial improvements in resolution, only SMLM techniques allow for resolution on the scale needed to describe spatial relationships in the tens of nanometers.

quisitions, reconstruction artefacts have been a concern. Acquiring an experimental point spread function (PSF) substantially reduces the likelihood of artefacts, as does the reduction of sample drift and chromatic aberrations[173]. Artefacts become more likely as a sample photobleaches and the signal-to-noise ratio (SNR) deteriorates - this is an important consideration in live cell imaging experiments where later acquisitions are likely to suffer from this. Similarly sample depth and restricting interference from out of focus light are critical to high quality SIM imaging[173].

Recent advances in the application of non-linear SIM have improved on the technique's 2-fold increase in lateral resolution. Non-linearity is easily introduced in samples with photoswitchable fluorophores and dyes, as well as simply through applying high illumination intensities to drive singlet state saturation[178, 179]. The development of photoswitchable fluorophores with high brightness and the ability to undergo a high number of switching cycles is extremely promising - particularly in the context of non-linear 3D SIM in lightsheet imaging (discussed in section 2.1.4)[180].

The relative ease of application and live cell utility of SIM have driven its popularity in recent years, particularly in the field of megakaryocyte and platelet biology. SIM has recently been proposed as a method of diagnosing platelet granule disorders by Westmoreland *et al.* as the two fold improvement in resolution allows for the accurate quantification of granule number [181]. Poulter *et al.* applied SIM to investigate the distribution and architecture of actin nodules in spreading platelets. Finally, Kahr *et al.* utilise SIM to investigate the effects of an ARPC1A knock-out in an immortalised megakaryocyte cell line[163].

2.1.2 STimulated Emission and Depletion Microscopy (STED)

Stimulated Emission and Depletion Microscopy (STED) refers to one of the original methods of SR imaging. STED is similar to confocal in that a pinhole is used to restrict illumination to a single image plane and achieve pixel by pixel excitation, however to achieve super-resolution,

the effective PSF is improved by applying a second laser (a depletion beam) with a red-shifted wavelength[182, 183, 184].This depletion laser is described as a doughnut-shape with 0 intensity at the centre to selectively suppress excitation around the centre of the imaging laser's PSF, thereby improving the effective resolution of the system[182].

An extension of this approach is Reversible Saturable Optical Fluorescence Transitions (RESOLFT) microscopy, where the spatial manipulation of both organic fluorophores and photoswitchable fluorophores into their respective on-off states can achieve improvements in lateral resolution[185].

Like SIM, STED can be performed on conventional fluorophores, and therefore offers relative ease of use in terms of sample preparation and in the absence of post-acquisition processing. As in any imaging modality, the quality of fluorescent labels is key to sample quality. Criteria like molecular brightness and photostability for example, will determine the quality of imaging. Critically, fluorophores with excitation/emission spectra and singlet states favourable for this modality must be chosen. Most of the complexity in STED arises from the illumination of the sample, in particular selecting appropriate excitation and depletion wavelengths. The magnitude of the depletion laser also restricts the application of STED based techniques to live cell imaging. Although RESOLFT in particular has been reported in the live cell format, ultimately this approach depends on significant intensities which will undoubtedly have phototoxic effects on cells.

A very recent and exciting application of STED has been the combination of expansion microscopy and STED imaging. This method, referred to as ExSTED, reportedly achieves a 30 fold increase in lateral (to below 10nm), and 50nm isotropic resolution[186].

2.1.3 Computational methods of super-resolution

In the last decade methods of widefield deconvolution have also offered a means of sub-diffraction limited imaging. Prominent examples include SOFI (Super-Resolution Fluctuation

Imaging), an approach which relies on the statistical analysis of fluctuations in fluorescent signal over time to achieve up to a five-fold improvement in resolution [187]. A Bayesian localisation approach developed by Cox *et al.* in 2011 achieved lateral resolutions of 50nm by modelling blinking and bleaching events in an entire dataset as opposed to the frame by frame approach typically applied in localisation microscopy[188]. These approaches can successfully extract super-resolution data from non-linear data sets, without the need for specialist optics (although they benefit substantially from high quality cameras and objectives).

Most recently, the SRRF method has offered a means of super-resolution through sub-pixel radial fluctuations, without the traditional computational cost associated with high density analysis[189]. This approach has substantial benefits in live cell imaging, as the authors demonstrate that low levels of illumination can be used to generate super-resolved images at sub-second temporal resolutions.

2.1.4 Super-resolution in lightsheet imaging

Selective plane illumination microscopy (SPIM), also known as lightsheet imaging, has become one of the most popular live cell imaging approaches in recent years. It is also used extensively in imaging larger fixed samples, including cleared bone and brain samples[190, 191]. The main advantages of SPIM are superior optical sectioning (due to the illumination of a single plane through the orthogonal arrangement of both illumination and detection pathways) and speed, both of which are critical in live cell imaging, where selective illumination substantially reduces the phototoxic effect of repeated illumination[190, 191]. As the focus of this chapter is super-resolution imaging, this section will focus primarily on lightsheet applications which have achieved sub-diffraction limited imaging.

The high SNR and rapid volumetric imaging offered by SPIM systems makes them ideal modalities for super-resolved imaging. This is particularly significant for live cell imaging due to the substantial reduction in phototoxicity achieved through selective plane illumination. A

number of approaches combining structured illumination with SPIM have successfully achieved live cell super-resolution imaging. Most notably, Li *et al.* demonstrate remarkable live cell sub-diffraction imaging through non-linear excitation in a lattice lightsheet set up[192, 193]. Chang *et al.* used two counter-propagating sheets to produce structural illumination (a method described as coherent structured illumination light sheet based fluorescence microscopy - csiLSFM)[180].

More recently single molecule localisation microscopy has been performed in lightsheet modalities. Galland *et al.* apply a method of SPIM referred to as soSPIM to perform single molecule localisation microscopy in 3D samples (cells and Drosophila), performing remarkably effective single molecule imaging at unprecedented depths in large samples. The use of tilted lightsheets to overcome the geometrical limitations of orthogonal excitation has allowed high numerical objectives to be used in SPIM set ups, with significant consequences for the application of SPIM for super-resolved imaging[194, 195]. Gustavsson *et al.* have demonstrated high quality 3D single molecule localisation microscopy using a tilted sytem. A recent application of the lattice lightsheet demonstrated 3D single molecule imaging at depth in *Drosophila* [196]. A particularly recent use of the lightsheet approach compared a traditional TIRF single molecule live cell acquisition on coverglass with lightsheet imaging of suspended T-cells[196]. The authors found that while the two imaging modalities were comparable in performance, receptor diffusion was profoundly affected by the presence of glass, and that contact with a surface resulted in calcium influx. These findings suggest that by removing the restriction of imaging at a coverslip, lightsheet imaging can also remove biological artefacts which arise as a result of technological limitations.

2.1.5 Single Molecule Localisation Microscopy (SMLM)

Of the many sub-diffraction limited techniques available, photoactivated localisation microscopy (PALM), and stochastic optical reconstruction microscopy (STORM) allow for the highest spatial resolutions[197, 198, 199, 200, 184]. These technologies can achieve localisation precisions up to 10nm, effective spatial resolutions of 20nm and below, and have also been shown to offer 3 dimensional, multi-colour, and live cell imaging options, all with distinctive benefits and drawbacks[197].

PALM and STORM are single molecule localisation microscopy (SMLM) techniques, reliant on the stochastic activation of subsets of molecules within a sample to allow for precise localisation and rendering post-acquisition[198, 199]. The key difference between these two approaches lies in the nature of the reporters used to generate 'blinking' events as a result of the stochastic activation of fluorophores (Figure 2.2). While in STORM paired dyes (nSTORM) or organic fluorophores capable of switching within an appropriate redox buffer (dSTORM) are used, PALM relies on fluorescent proteins with photoswitching capabilites which occur as a consequence of *cis-trans* isomerism when irradiated at particular wavelengths[197, 198, 199, 200, 184].

Each technique offers its own distinct disadvantages, dSTORM is limited to fixed samples and is dependent on the use of high quality antibodies, as well as the optimisation of imaging buffers. PALM suffers from a limited palette of photoswitchable tags, decreased photon counts and increased localisation uncertainties, as well as the artefacts associated with over-expression as tags often must be transfected into cells. Quantitatvely, dSTORM is often confounded by unknown labelling stoichiometries and the undefined on-off cycling of organic fluorophores.

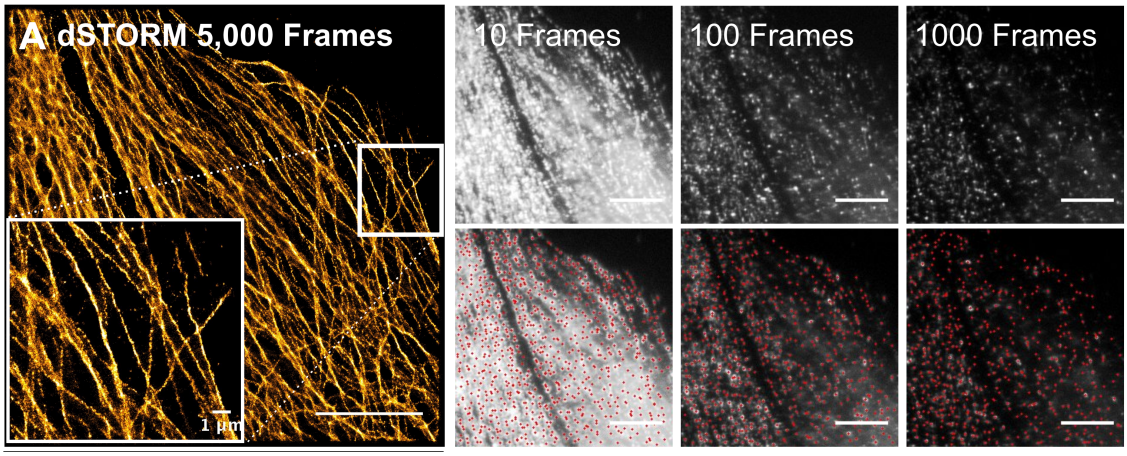

Figure 2.2: **Single molecule localisation microscopy (SMLM).** SMLM approaches involve generating a data set comprised of consecutively acquired frames capturing the activation of subsets of fluorophores. Provided sufficient temporal separation is achieved by these acquisitions, each individual emission can be fitted to a Gaussian to determine its centroid, and therefore reconstruct a super-resolved image from the detection and recording of potentially millions of single localisation events as samples progressively bleach.

Both dSTORM and PALM have been applied successfully in a number of studies demonstrating not only the utility of nanoscopic resolution, but the potential of quantitative single molecule microscopy. In a recent review by Miklosi *et al.*, the role of single molecule imaging in dissecting the function and distribution of neurotransmitters and other molecules of interest at synapses is discussed in the context of quantitative proteomics [201]. The authors discuss the utility of dSTORM as a quantitative and spatial tool in neurobiology and explore studies where the application of single molecule imaging allows for the mapping and quantification of vesicles and neurotransmitters in synapses.

In platelet biology recent work by Mayr *et al.* demonstrate super-resolution of the platelet cytoskeleton [202]. Poulter *et al.* applied dSTORM as a quantitative tool to interrogate integrin distribution at actin nodules and show a distinct absence of integrin clustering at nodules - a finding which would not have been possible without the use of single molecule localisation microscopy[27].

Both PALM and dSTORM are potentially potent quantitative tools, however single molecule imaging and quantification have important caveats which are critical for accurate image acquisition, reconstruction, and quantification. First and foremost is the correct identification of single molecules, which is often confounded by frame to frame persistence. The phenomenon of 'blinking' involves repeated cycling between on-off states, and persistence across multiple frames such that 'single' molecules are more often collections of peaks [203, 204, 205]. This is a significant caveat in single molecule quantification, and a considerable focus of the field is accurately detecting individual emitters and modelling fluorophore behaviour.

A significant advance in the field would be the development of systems with an established labelling stoichiometry and known photoswitching kinetics. While PALM addresses both to an extent by introducing photoswitchable tags with a known 1:1 labelling ratio, the variability and artefacts introduced by over-expression are one of a number of problems inherent to

this approach. As the following section will discuss, the use of CRISPR to knock-in tags to ameliorate the negative effects of PALM has the potential to substantially advance single molecule imaging.

2.1.6 Development of CRISPR-PALM

While there have been reports of successful CRISPR mediated knock-in of fluorescent proteins since the development of the tool, the generation of endogenously expressed fusion proteins remains a relatively new field with limited applications. To date, only one study has reported the use of knock-in reverse switchable mEGFP to generate super-resolution data through the RESOLFT method[206]. A recent study used a SMLM approach to investigate RNA Polymerase function through knock-in of the photoswitchable tag Dendra - however the authors do not adequately report super-resolution or investigate the effect of knock-in on protein expression and function. Perhaps most significantly, this study does not compare the resulting data with any other methods of super-resolution, raising doubts as to the efficacy of their super-resolution method[207].

The use of endogenously expressed fluorescent proteins would address a number of long standing issues in single molecule imaging. Most significantly, a known, defined ratio of labelled protein would substantially improve the quality of single molecule imaging. By avoiding over-expression transfection artefacts like the artificial clustering of proteins, would be prevented. Finally in the age of SPIM and long term fluorescence imaging at low phototoxicities, endogenous expression offers the opportunity to observe cells over significantly longer periods than those typically defined by transfection.

The work reported in this chapter describes efforts to develop CRISPR-PALM and address some of the obstacles presented by endogenous expression. While CRISPR knock-in provides a means by which to avoid over-expression and its associated issues, a key question is whether endogenous expression can achieve the labelling density required to satisfy the Nyquist-

Shannon criteria. Key aims of this section include first establishing methods by which to introduce fluorescent tags to target loci through CRISPR knock-in, and to establish whether the resulting expression levels are sufficient for single molecule imaging.

The work in this chapter focusses on the α-tubulin isoform expressed by the gene *TubA1B*. Firstly, tubulin is in many ways a perfect target for single molecule imaging as microtubules are found well beneath the diffraction limit, with an average diameter of approximately 25nm. With monomers 4nm apart from one another, resolving these fine sub-cellular structures will require an even distribution of reporter fluorophore at a high enough density to satisfy the sampling criteria. Secondly, this gene is ubiquitously expressed and as such will allow us to interrogate any cell line specific effects (as different lines will express different amounts of this specific isoform), and so establish where the labelling criteria for SMLM are met. Finally, as this gene is expressed, single cell sorting to isolate clones and to measure knock-in efficiency will be valuable tools in establishing CRISPR-knock in methodologies in this project.

As our long term interest is to study the molecular causes of inherited thrombocytopenia, we will adapt these methods to study the megakaryocyte and platelet specific β-tubulin (detailed in Chapter 4), the work performed in this section is relevant to adapting CRISPR-knock in to iPSCs as discussed in chapter 3.

2.2 Materials and Methods

2.2.1 Cell culture and transfection

Human Alveolar Basal Carcinoma (A549) and Embryonic Kidney 293T (HEK293T) cell lines were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with Penicillin/Streptomycin (1%), 2mM L-Glutamine, and 10% Foetal Bovine Serum (FBS). Human Erythroleukaemia 92.1.7 cells (Hel 92.1.7) were maintained in identically supplemented Roswell Park Memorial Institute (RPMI) 1640 medium. All cells were cultured at 37°C and 5% CO_2. Adherent cells were passaged by washing T75 cell culture flasks (Corning) twice with sterile Phosphate Buffered Saline (PBS) without calcium or magnesium, before incubation with Trypsin-EDTA (Thermo Scientific) for 5 minutes at 37°C. Cells were then re-suspended in complete DMEM before passaging at an appropriate dilution in a fresh flask. Suspension cells (Hel 92.1.7) were passaged by spinning cells at 1200 RPM for five minutes before re-suspension in fresh complete RPMI, dilution in fresh media, and incubation.

Adherent cells were transfected with Lipofectamine 3000 (Thermo Scientific) as per manufacturer guidelines. Briefly, cells were plated at a density of $1x10^5$ in 12 well plates (Thermo Scientific) 18 hours before transfection. Immediately preceding transfection, cells were washed twice with sterile PBS before incubation in Optimem. For Dronpa experiments, the BPK1520 guide expressing vector (Addgene Plasmid #65777 BPK1520 was a gift from Keith Joung), Cas9-RFP (Sigma), and Dronpa donor vectors were transfected at equimolar ratios to a total of 1μg DNA per well. For mEGFP and mEos experiments, the px459 guide and Cas9-puro expressing plasmid was used (Addgene plasmid #62988 was a gift from Feng Zheng).

Suspension cells were transfected using the Neon electroporation system (Thermo). Briefly, cells were resuspended in a final total volume of 10μl buffer R with 1μg of total DNA (equimolar guide C px459 and donor) and a total of 1 x 10^5 cells. 10 μL tips were used with 2x pulses at

20ms pulse width and 1450V.

2.2.2 Guide vector cloning

The MIT CRISPR database (http://crispr.mit.edu/) was used to design guide sequences, each of which were ordered as complementary oligos and suspended at 10μM before incubation in the presence of T4 PNK (New England Biolabs) to generate phosphorylated oligos. Annealing was performed using a thermal cycler (Sense Quest) programmed to a 5 minute 95°C denaturing cycle, followed by a gradual cooling to room temperature at - 0.1°C per second. The BPK150 vector was digested with the BsmBI restriction enzyme (New England Biolabs) and treated with T4 Calf Intestinal Phosphatase (CIP) (New England Biolabs) to generate linearised, dephosphorylated backbones with overhanging sequence complementary to the annealed oligo duplex.

Finally, both digested vector and oligo duplex were ligated using T4 DNA ligase (New England Biolabs) at room temperature for 10 minutes. 2μl of the resulting ligated DNA was then transformed into competent cells (Bioline - Alpha Select Silver Efficiency *E. Coli*) by heat shock at 42°C for 30 seconds, incubation on ice for 2 minutes, and incubation in SOC medium (Thermo) for 1 hour. Transformed cells were plated on penicillin agar plates over night before colony selection and screening.

Colony PCRs were used to validate guide vectors. 2μL of each bacterial inoculation was amplified using an OS280 primer and RedTaq ReadyMix PCR Reaction mix. Thermal cycling was performed according to manufacturers instructions, and the resulting product run on a 1.5% agarose gel with ethidium bromide for 25 minutes at 120V.

2.2.3 Donor vector cloning

Donor plasmids were generated by ordering gBlock synthetic DNA (Integrated DNA Technologies) for both left and right homology arms and each of the inserts tested (Dronpa, mEos

3.2, mEos 3.2 codon optimised, mEos 4b, mEos 4b codon optimised, and mEGFP). Each fragment was designed with a 20bp segment overlapping the adjacent arm. Homology arms and inserts were cloned into an empty pGem-T-Easy backbone using a HiFi Gibson Assembly kit (New England Biolabs).

2μL of each 10μL reaction volume was transformed into competent cells as described in the previous section. Individual colonies were selected after overnight growth on ampicillin plates, miniprepped (using manufacturer instructions included in the GeneJet kit - Thermo Fisher), and subject to an EcoRI test digest to verify the presence of an insert. Clones carrying an insert of the correct size were further verified by Sanger sequencing.

2.2.4 Flow cytometry

Validation and verification of knock-ins was performed using an Accuri C6 Flow Cytometer (BD Technologies). Samples were gated according to forward and side scatter, with positive fluorescence for these experiments detected in the FL-1 channel.

2.2.5 Single cell sorting

Single cell sorting was very generously performed by Matt MacKenzie (TechHub) on a BD FACSAria Fusion cell sorter with a 100μm nozzle at 20 psi. Gates were set on the brightest population of cells to ensure that both highly expressing and correctly editing clones were selected.

2.2.6 Western Blotting

Clonal populations were expanded in 6-well plates (Thermo Scientific) before lysis in NP-40 buffer with proteolysis inhibitor (Sigma). Lysates were prepared by 30 minutes incubation on ice in NP-40 with inhibitors before a final 10 minute spin at 14000 rcf. The supernatant was then added to 2x reducing sample buffer and boiled for 5 minutes.

Western Blots were prepared through SDS-PAGE on 4-12% gradient Bolt gels (Thermo Scientific). Gels were run for 15 minute at 70 V, and 45 minutes at 125 V. Once run, each gel was transferred to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a Turbo transfer system (Bio-Rad).

After transfer, membranes were blocked with 4% BSA in 0.1% Tween-20 Tris buffered saline (TBST) and probed with the relevant primary antibodies. Antibodies to α-tubulin (Sigma - T9026), GAPDH (Abcam) and GFP (Thermo Scientific) were predominantly used in this chapter. To determine cell line variability in the expression of *TubA1B*, a *TubA1B* specific antibody (Abcam) was used. Values were normalized to loading controls (GAPDH). Secondary incubations were performed using fluorescent antibodies (LiCor Instruments) in TBST. Anti-mouse 680 and anti-rabbit 800 fluorescent antibodies were used for detection using an Odyssey Fc (LiCor Instruments). Finally, Image Studio was used to quantify western blots.

2.2.7 Imaging (PALM, dSTORM, and SRRF)

Clones of interest were imaged on 35mm MatTek Dishes (MatTek Corporation) on a Nikon N-STORM system (Andor iXon Ultra DU897U EMCCD, Ti-E stand, Perfect Focus, Agilent MLC400 laser bed). A 100x 1.49 NA TIRF Objective was used for each acquisition, which was then reconstructed using ThunderSTORM [208] (Maximum Likelihood, Integrated Gaussian PSF fitting). Live cell SRRF images were reconstructed from 100 frame acquisitions using NanoJ-SRRF[189].

Hel 92.1.7 cells were treated with phorbol 12-myristate 13-acetate (PMA) (Sigma) and thrombopoeitin (TPO - a generous gift from Ian Hitchcock) overnight to drive spreading and differentiation prior to imaging. Once seeded on MatTek dishes, cells were washed twice with PBS before treatment with microtubule stabilising buffer (MTSB - 80mM PIPES pH 6.8, 1mM $MgCL_2$, 4mM EGTA) and 0.5% Triton-X100 for 30 seconds before fixation with ice

cold Methanol at -20°C for 3 minutes. Samples were then washed with TBST and imaged in PBS for CRISPR-PALM. dSTORM samples were blocked (2% goat serum,1% BSA) before staining with α-tubulin primary, and anti-mouse Alexa-647 secondary. Finally, these samples were imaged in blinking buffer (100mM mercaptoehylamine-HCL, 1g/mL catalase, 50g/mL glucose oxidase, PBS).

2.2.8 Statistical Analysis

Statistical analysis was performed using GraphPad PRISM 6, with statistical tests as indicated in figure legends. Briefly, significance was determined using a 2-way ANOVA with multiple comparisons (Tukeys). $n = 3$ across samples, and error bars represent standard deviation of the mean.

Full Width Half Maxima (FWHM) were measured across microtubules by drawing a 5μm line on ThunderSTORM reconstructed normalised Gaussian images. Each plot profile was uploaded to Igor Pro (6.37) for multi-peak fitting. Fourier Ring Correlation for mEos 3.2 and dSTORM data was performed using the BIOP FRC plug-in. For mEos 4b data this was performed using the recently published NanoJ-SQUIRREL plug-in[209, 210].

2.3 Results

To date, no other attempts at generating super-resolution single molecule data sets through CRISPR knock-in have been reported. The closest study describes the insertion of Dendra into RNA Polymerase II, but the authors do not evaluate the effects of the knock-in on expression, nor on the resolution of the final image. The stringent sampling requirements of imaging at 10-20nm resolution is critical in experiments where label density will be determined by genomic expression. In classical PALM the use of over-expression, while biologically problematic (particularly in receptor quantification), is helpful in that it provides an abundance of labelled protein - and so satisfies this otherwise strict sampling criteria.

To that end the results that follow work to first establish the validity of using photoswitchable tags for SMLM through endogenously expressed reporters.

2.3.1 CRISPR-Cas9 mediated knock-in of the photoswitchable tag Dronpa into the *TubA1B* locus

While our long term goal is to study the effects of patient mutations on *TubB1* and platelet production, the absence of proplatelet forming cell lines means that the only viable method of phenotypically screening the effect of mutations would be the use of iPSC differentiated MKs. Unfortunately the literature to date shows that HDR efficiency is extremely low in iPSC, making genome editing in iPSC a poor choice for developing and validating a methodology. Similarly, at the beginning of this project we had no protocols for the differentiation of iPSCs into mature, proplatelet forming MKs. As *TubB1* is only expressed in terminally differentiated megakaryocytes, this would hinder our efforts to investigate the feasibility of using genomically introduced tags for our imaging.

With this in mind, our initial experiments targeted a ubiquitously expressed tubulin gene, *TubA1B* in Hek293T. These cells have been extensively used in the literature to date as they demonstrate a reasonably high rate of HDR and knock-out. We selected the photoswtichable

tag Dronpa for our pilot experiments, Dronpa is a monomeric fluorescent protein with high brightness and has been favourably reported in its applications in PALM[211, 212].

4 guide sequences were designed using MIT's Zhang lab guide tool (described in methods), each targeting the N-terminus of the *TubA1B* gene (Figure 2.3A). Oligonucleotides for each guide were synthesised, annealed, and cloned into BPK1520 empty vectors to form *TubA1B* targeting guide vectors. These were validated first using a colony PCR, and finally by sequencing (Figure 2.3B).

An HDR template was generated by designing homology arms flanking the N-terminus of *TubA1B*, between which a Dronpa sequence was inserted. Each segment was ordered as a gBlock from IDT and assembled into a pGem-T-Easy backbone (Figure 2.3C). While the original homology arms were designed to be approximately 600bp each, issues in the synthesis of the left (5') arm resulted in a final design of 220bp for the left homology arm, 690bp for the Dronpa insert, and 626bp for the right homology arm (Figure 2.3D). The presence of an insert was first verified with an EcoRI restriction digest, with successfully cloned plasmids appearing 1536bp in size compared to the 3000bp of the empty pGem-T-Easy backbone (Figure 2.3E).

Representative control transfection FACS plots are demonstrated in figure 2.4A, where no fluorescence is observed in the FL-1 channel in both untreated and Cas9-RFP only controls. Conversely, 90% of cells transfected with a Dronpa α-tubulin over-expression vector are positive in this channel. In cells treated with a mix of Cas9-RFP, donor vector, and the relevant guide a small percentage of cells become positive in the FL-1 channel (Figure 2.4B) suggesting successful HDR and integration of the donor repair template.

A comparison of all 4 guide sequences shows that guide C is the most efficient at facilitating the insertion of the Dronpa tag at the target sequence (Figure 2.5). No significant difference in median fluorescence intensity (MFI) is evident across any of the guides, consistent with the fact that the introduction of the sequence for endogenous expression is unlikely to yield

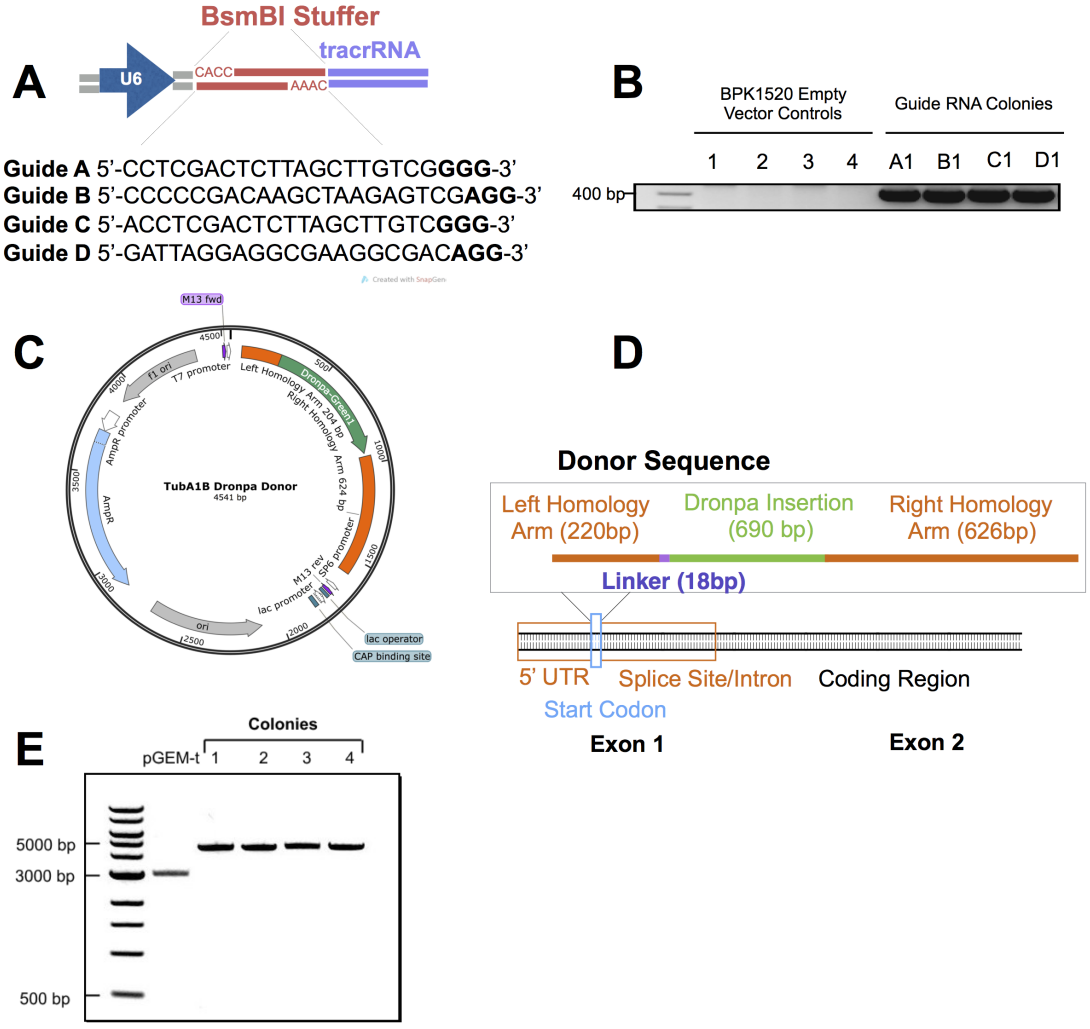

Figure 2.3: **Generation of guide and donor plasmids for CRISPR knock-in targeting the *TubA1B* locus.** (A) Guide RNAs targeting the N-terminus of *TubA1B* were cloned into a BsmBI stuffer cassette in a BPK1520 backbone. (B) Successful cloning was validated using a colony PCR which successfully amplified the relevant sequence in each cloned plasmid. (C) To introduce the Dronpa tag into the targeted locus, an HDR template with the fluorescent protein flanked by homology arms was designed and cloned into a p-Gem-T-Easy backbone. (D) This donor plamid featured 3000 base pairs (bp) of backbone, a 690bp Dronpa sequence, and left and right homology arms (220 and 626bp respectively). (E) Successfully cloned donor plasmids were validated by restriction digest before sequencing. Control pGem appeared as a linearised 3000bp sequence, while successfully cloned HDR template plasmids appeared as approximately 1500bp heavier.

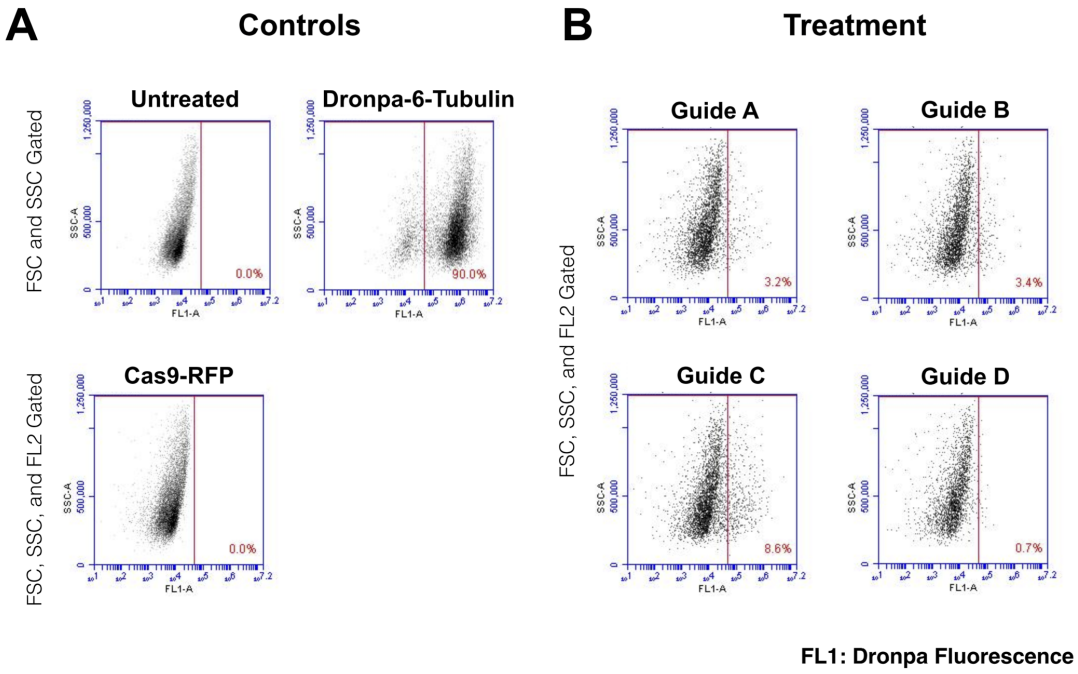

Figure 2.4: **Successful integration of fluorescent donor template in cells transfected with guide RNA and donor plasmid.** Hek293T were co-transfected with Cas9-RFP, guide, and donor plasmids. (A) Untreated and Cas9-RFP only controls show no fluorescence in the FL1 channel, while over-expressed Dronpa α-tubulin results in a strong shift. (B) In cells co-transfected with Cas9, guide, and donor plasmids a percentage of cells becomes FL-1 positive, indicative of successful HDR.

any variation. Interestingly however, the MFI of each population is dramatically lower than the over-expression control.

The highest efficiency guide (guide C) was taken forward to generate clonal populations expressing tagged *TubA1B*. After single cell sorting, clones were expanded before DNA extraction and subsequent genotyping. Primers designed as indicated in Figure 2.6 amplified a 324bp wild type sequence, and where the Dronpa tag had been successfully inserted, a 1018bp fragment. The clones selected demonstrate a mix of different genotypes including hetero- and homo-zygous cells. This result shows a successful insertion, and so the brightest of these cells as determined by FACS were taken forward for further evaluation.

Dronpa clones 1, 4, 6, 7, and 11 were taken forward for further evaluation. Firstly, cells were tested by flow cytometry after PCR validation to ensure fluorescent protein (as a genomically expressed *TubA1B* fusion) was expressed. As shown in figure 2.7, each clone was notably fluorescent in the FL-1 channel when compared to wild type. Lysates were then collected for each clone and western blotting was performed to determine whether this fluorescence was due to the expression of a Dronpa tagged α-tubulin.

Lysates were prepared for wild type Hek293T, validated Dronpa positive CRISPR clones, and Hek293T transfected with Cas9-RFP and donor only. In WT, Cas9, and donor transfected samples a mouse monoclonal antibody to α-tubulin detects a single band at the predicted molecular weight of approximately 50 kDa (Figure 2.8A) cells transfected with 100ng and 500ng of the Dronpa over-expression plasmid, a second band at a molecular weight of 80 kDa is observed. This is consistent with the predicted molecular weight of an α-tubulin fusion with Dronpa. Each of our CRISPR clones are positive for a band at this weight, indicating that a fusion protein of the correct molecular weight is expressed by each of the cells (Figure 2.8A). This was further confirmed by probing with a Dronpa specific monoclonal antibody which recognises this 80kDa band(Figure 2.8B).

Surprisingly there is no difference in expression between heterozygous (clones 1, 4, 11) and

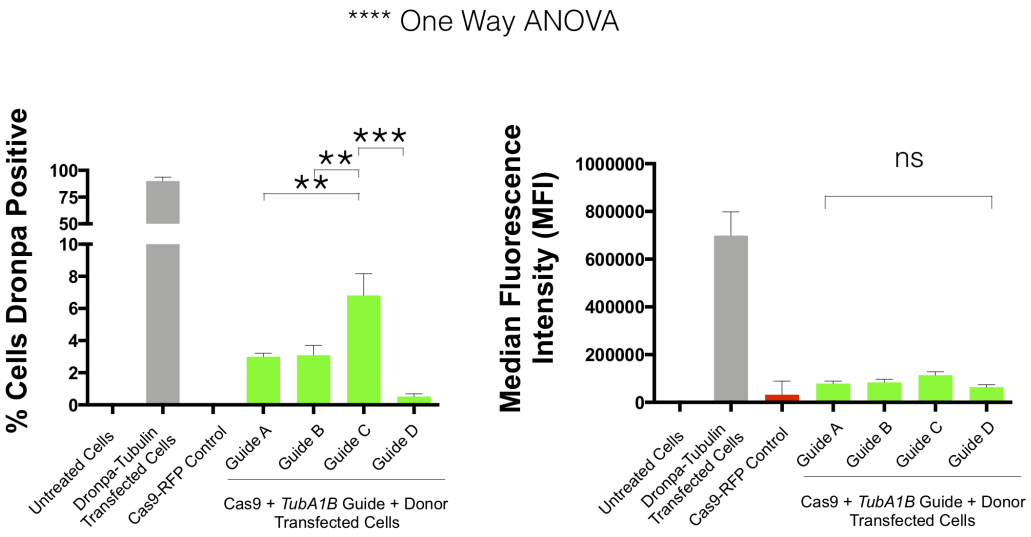

Figure 2.5: **Efficient knock-in using *TubA1B* targeting guide C.** Hek293T cells were co-transfected with Cas9-RFP, donor, and guide plasmids. Compared to untreated and Cas9 only controls, cells transfected with all 3 plasmids demonstrate a percentage of FL-1 positive cells. Of the 4 guides tested, guide C presents the highest rate of HDR integration of the Dronpa tag (6.82% ±1.2). No significant difference in median fluorescence intensity is evident across any of the guides. One-way ANOVA with multiple-comparisons performed. $n = 3$.

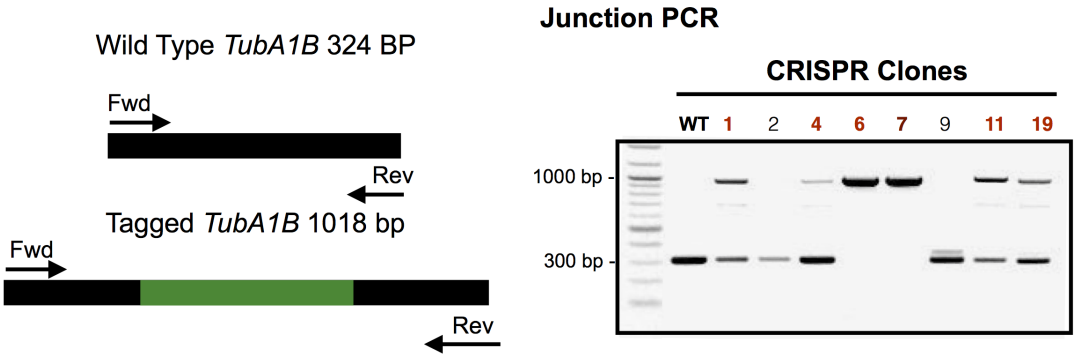

Figure 2.6: **PCR validation shows successful knock-in at target locus.** After generating knock-in *TubA1B* Hek293T cells, individual clones were isolated using fluorescence assisted cell sorting (FACS). Each clone was expanded before DNA extraction and genotyping by PCR. Primers designed to flank the insert site generate a 324bp band for wild type sequence, and a 1018bp band for successfully inserted alleles. Compared to WT cells, individual clones demonstrate a mix of heterozygous and homozygous insertions of the Dronpa sequence.

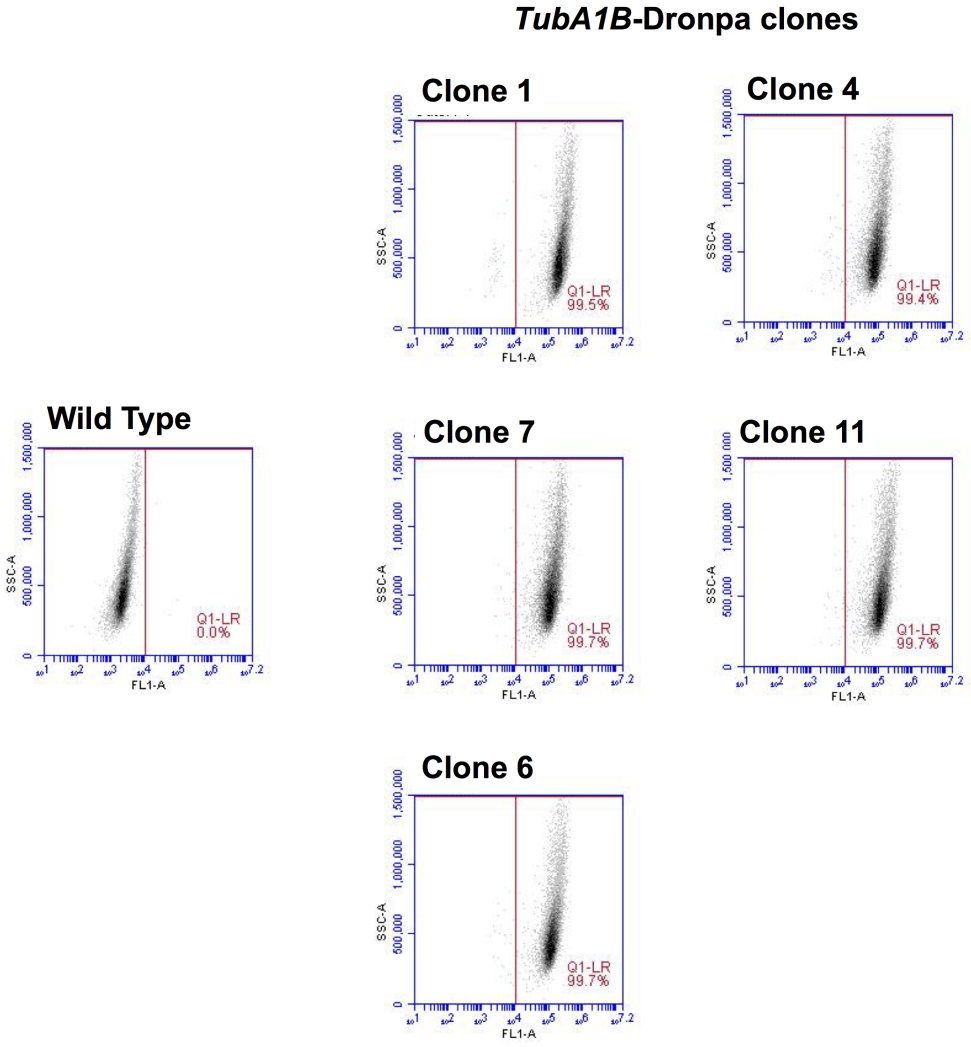

Figure 2.7: **FACS of individual clones shows the presence of a fluorescent population.** Hek293T *TubA1B*-Dronpa clones validated by PCR were screened by flow cytometry to ensure that a fluorescent protein was expressed. All 5 clones are positive in the FL-1 channel when compared to wild type Hek293T, indicating that a functional fluorescent protein is expressed in these cells.

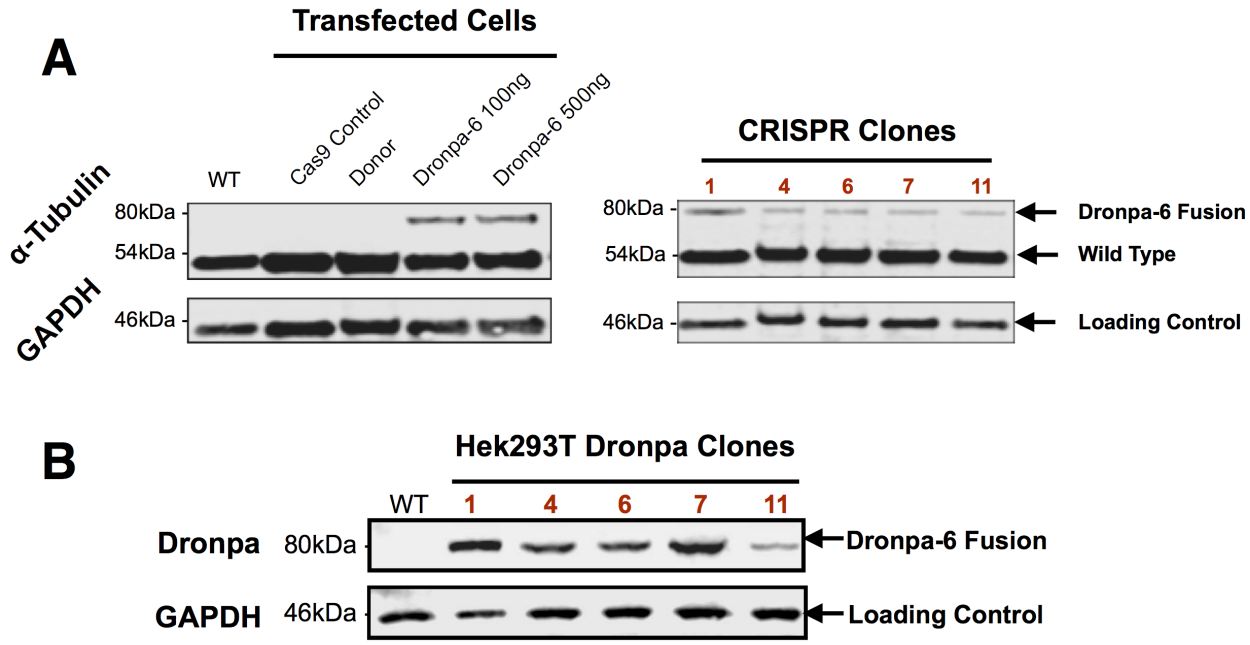

Figure 2.8: **Western blotting of Dronpa tagged *TubA1B* Hek293T indicates expression of tagged tubulin.** CRISPR tagged Hek293T were single cell sorted and expanded for western blot analysis. (A) Negative control cells include wild type, untreated Hek293T, as well as cells transfected with Cas9 vector (no guide or donor), and a donor only control. In these samples, α-tubulin appears as a single band at approximately 50 kDa. As a positive control, Hek293T were transfected with two concentrations of an over-expression Dronpa tubulin vector, and in these samples a second band appears at approximately 80 kDa, the predicted weight of a Dronpa α-tubulin fusion protein. In each of the PCR verified CRISPR clones (1, 4, 6, 7, and 11) this heavier band is evident. (B) Lysates were probed with a Dronpa specific antibody, and this detects the same band confirming the presence of a Dronpa labelled α-tubulin in each CRISPR clone.

homozygous clones (6 and 7). The loss of wild type *TubA1B* would in theory result in a loss of the wild type protein, and as a result an increase in the expression of the Dronpa tagged protein when compared to heterozygous clones. The fact that this is not observed suggests that there may be a regulatory mechanism which silences a tagged gene if there is a negative or deleterious effect on protein folding or function. Finally, there is no significant loss of wild-type α-tubulin in homozogous clones suggesting that expression of the tagged allele is regulated by compensatory mechanisms potentially increasing the expression of other WT α-tubulin isoforms. Alternative explanations include the potential presence of an additional copy of the wild type protein, as well as potential translocations or deletions in one of the targeted alleles which results in an erroneous insertion at one of the targeted alleles.

2.3.2 Evaluation of endogenously expressed Dronpa tagged α-tubulin as a method of single molecule imaging.

To investigate whether the Dronpa fusion maintains the photoswitchable characteristics of the fluorophore, PALM imaging was performed on clone 7 and compared to dSTORM imaging of α-tubulin in Hek293T (Figure 2.9). Unfortunately the resulting Dronpa image appears poor both in widefield and as a super-resolved image, with indistinct microtubules when compared to the dSTORM equivalent. This was verified by comparing the full width half maxima (FWHM) of dSTORM with that of the Dronpa images (Figure 2.10). While distinctive peaks with FWHM ranging as low as 41.96nm are apparent in dSTORM image plots, the line profile from the Dronpa-PALM image is poor with FWHM as high as 232.06nm.

SMLM data sets can be compared quantitatively by looking at the total number of detections or localisations as in Figure 2.11A, which shows a significantly reduced number of detections in Dronpa Clone 7 when compared to the dSTORM image (** p =0.0016). Similarly, the mean xy uncertainty of the Dronpa cell is substantially higher than that of the dSTORM image, though this is unsurprising considering the dramatically higher photon

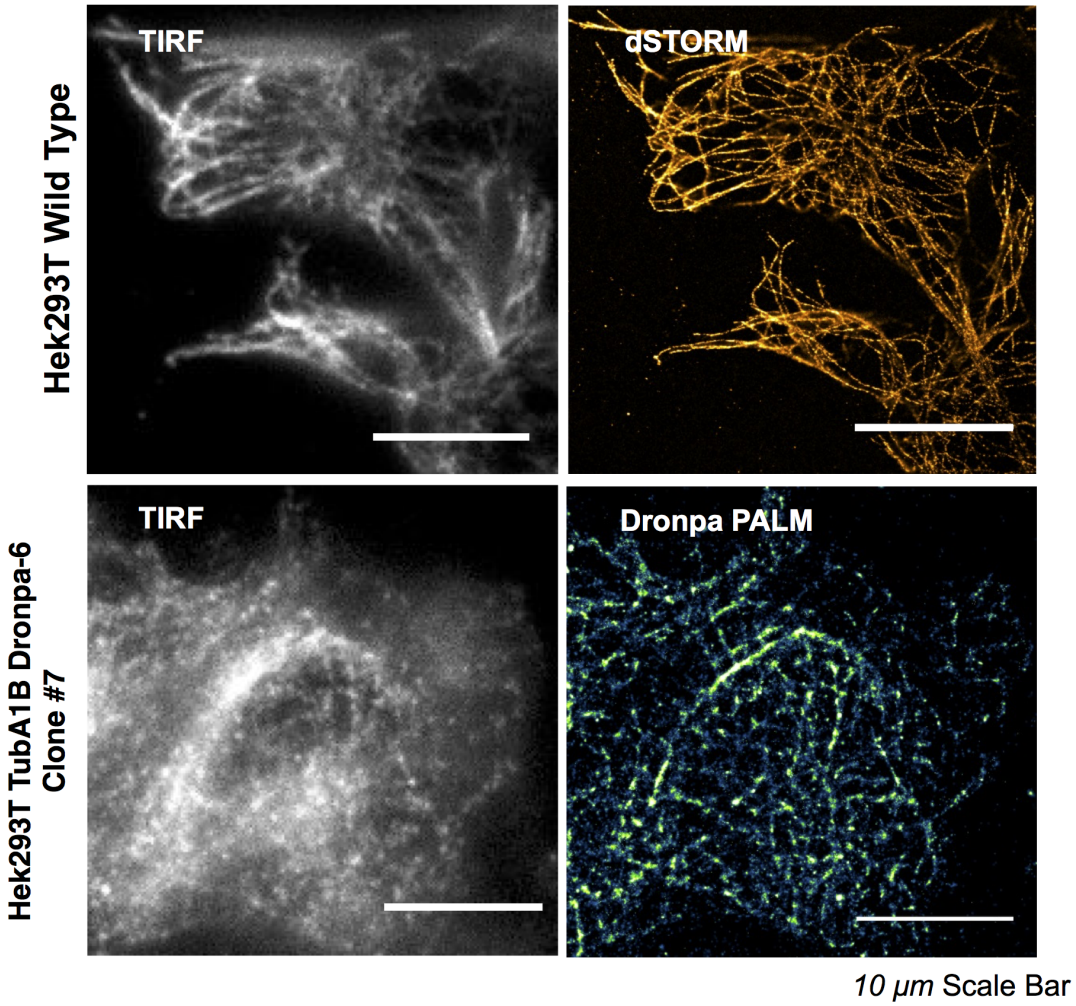

Figure 2.9: **Comparison of dSTORM and CRISPR-PALM Dronpa images shows poor quality SMLM in Dronpa knock-ins.** In Hek293T wild type cells labelled with an α-tubulin antibody, dSTORM imaging shows consistent microtubules. In comparison Dronpa PALM imaging of a CRISPR knock-in clone shows an uneven distribution of the labelled protein.

count of organic fluorophores (** p = 0.0191). Finally, as demonstrated by Figure 2.11B, the number of detections per frame rapidly tails off in the Dronpa clone, while it is sustained over the 20,000 frames observed in the dSTORM image.

Together this data shows that while a successful insertion of Dronpa at the target *TubA1B* locus successfully results in the genomic expression of a photoswitchable tag, there are a number of considerations which appear to impact the quality of the resulting SMLM image. The above comparisons to dSTORM, the gold standard of SMLM in terms of resolution, show that Dronpa tagging of the *TubA1B* gene does not produce high quality single molecule images. This is likely due to the low expression of the tagged protein across all Dronpa clones, regardless of whether they are hetero- or homo-zygous for insertion, as well as the poor performance of Dronpa as a SMLM reporter in the context of endogenous expression, despite previous applications in the literature in standard PALM.

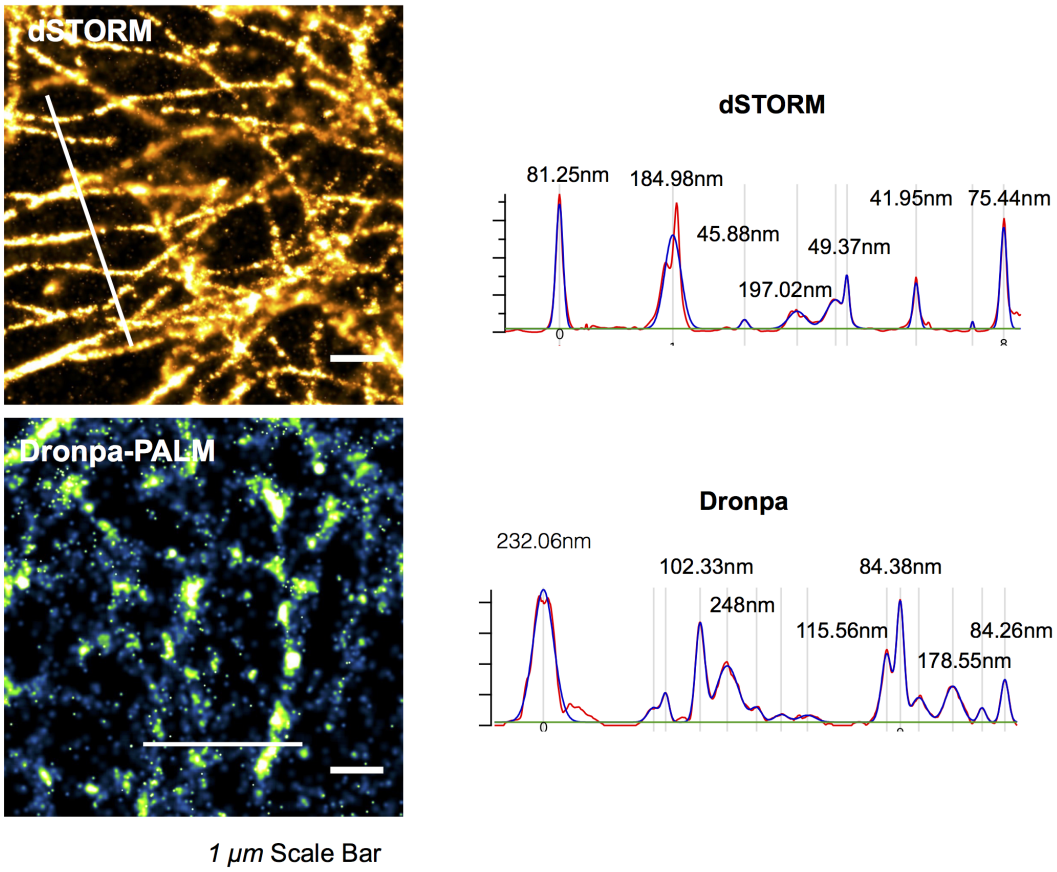

Figure 2.10: **Comparison of full width half maxima (FWHM) in dSTORM and CRISPR-PALM Dronpa images shows poor resolution of microtubules in Dronpa knock-in clones.** FWHM are a classical means of comparing effective resolution in super-resolved images. In these cropped images, there are substantial qualitative and quantitative differences in the performance of CRISPR-PALM when compared to dSTORM. While dSTORM shows continuous microtubules and bundles, the Dronpa image shows poor label distribution, and as a result poorer FWHM. Red lines indicate raw peak values as determined by line profiles in ImageJ, while red lines indicate the 'fit' and resulting FWHM as measured by automated peak fitting in IgorPro.

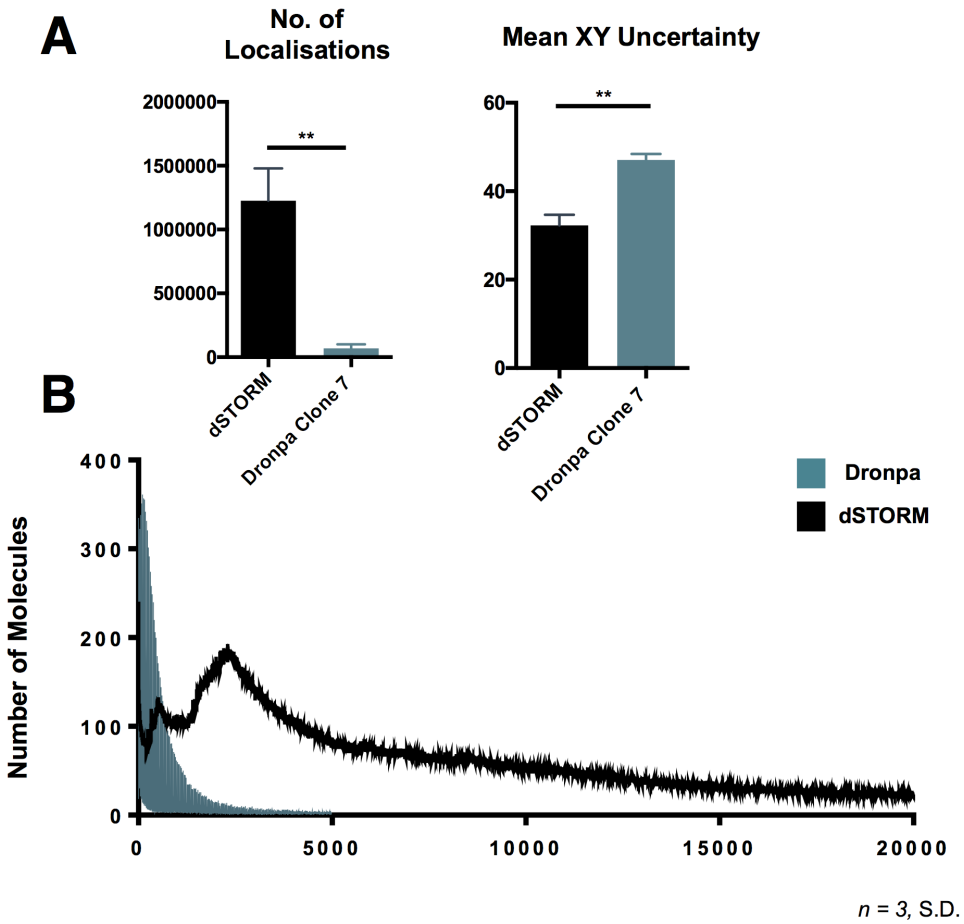

Figure 2.11: **Quantitative comparisons between dSTORM and CRISPR-PALM through Dronpa insertion support poor performance of Dronpa knock-in as a single molecule system.** Hek293T wild type cells imaged with dSTORM were compared to Dronpa tagged CRISPR clones. (A) Compared to dSTORM, Dronpa tagged clone 7 shows a significantly reduced number of localisations, consistent with the poor resolution in the images previously shown (** p =0.0016). Similarly Dronpa images show a significantly higher mean XY uncertainty indicative of poor SMLM imaging (** p = 0.0191). (B) A comparison of detections per frame shows that the Dronpa sample quickly tails off, while dSTORM imaged cells continue to detect molecules over the 20,000 frames observed. $n = 3$, S.D., Student's T-Test performed.

2.3.3 mEGFP and mEos 3.2 tagged *TubA1B*

The work presented thus far in this chapter demonstrates the successful insertion of a photoswitchable fluorescent tag in a targeted genomic locus through CRISPR-Cas9. However, these results show that the process of generating a robust method of single molecule imaging through genomic knock-in will have its own complications. Firstly, the expression level of the tagged allele observed is lower than expected, the fact that even homozygous knock-ins do not appear to surpass a particular level of expression suggests that there is a regulatory mechanism which maintains cell function and morphology at the cost of the expression of a targeted allele. This could be due to the tag used and the properties of both its sequence and encoded protein (monomericity, maturity, etc.), or due to the cell line used.

To address where and why Dronpa-*TubA1B* performed poorly as a method of CRISPR-PALM, donor vectors were redesigned to include mEGFP and mEos 3.2 sequences, flanked by homology arms and a linker identical to the Dronpa donor. mEGFP is a highly monomeric, rapidly maturing, codon optimised, bright fluorescent protein. mEos 3.2 is a photoconvertible fluorescent protein which shifts its emission spectra upon excitation with ultraviolet light[213, 214]. This photoconversion from an emission maxima of approximately 460 to 560 combined with subsequent bleaching at the 560 wavelength allows for the generation of stochastically activated fluorophores ideal for PALM. mEos 3.2 offers distinct advantages over other photoconvertible fluorophores - perhaps most significantly a higher photon count which is critical to high quality PALM[214]. While the fluorophore does not retain the brightness of Dronpa, nor its capacity for reversible photoswitching, it offers an important alternative in this work. Namely photoswitching kinetics which can improve the poor quality of CRISPR-PALM reported thus far.

Donor vectors were designed and cloned according to the same criteria as the previously reported Dronpa donor (Figure 2.12). Identical gBlocks for left and right homology arms were cloned into a pGem-T-Easy vector with either an mEos 3.2 or mEGFP donor (Figure

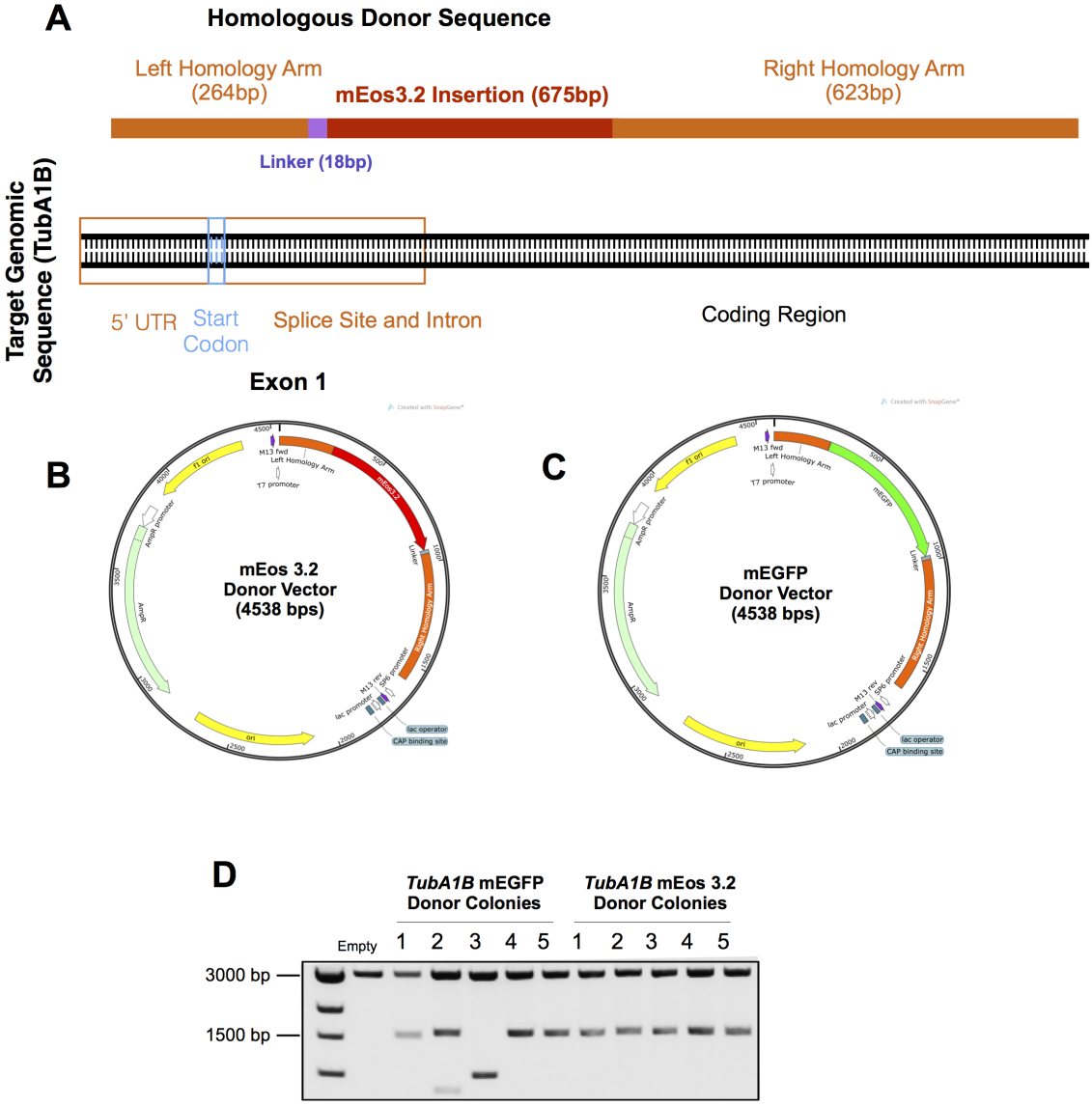

Figure 2.12: **Design and cloning of *TubA1B* mEGFP and mEos 3.2 donor vectors.** (A) mEos 3.2 and mEGFP inserts were designed using the same approach as the Dronpa approach, with the candidate insert flanked by homology arms and a short linker. (B) gBlocks for an mEos 3.2 insert were ordered with overlapping arms flanking the fluorophore for cloning into a pGem-T-Easy backbone. (C) A similar gBlock was synthesised to generate an mEGFP donor vector. (D) Inserts were cloned in using Gibson Assembly, with individual clones validated using an EcoRI digest. An EcoRI digested empty pGem-T-easy control vector runs at 3000bp, while a correct insert will result in two bands, one at 3000bp and the other at approximately 1500bp.

2.12A, B, C). Clones were subject to a test EcoRI digest and compared to a blank donor vector (Figure 2.12D). Clones positive for the correct insert presented with both a 3000bp and 1500bp band. Positive clones were sequenced to verify correct assemblies and taken forward for transfection experiments.

Hel 92.1.7, A549, and Hek293T cells were co-transfected with a px459 guide RNA expressing vector carrying the previously reported guide C and either the successfully cloned mEGFP or mEos 3.2 donor vectors. Cells were then single cell sorted and surviving clones expanded before DNA extraction and PCR to validate the presence of an insertion at the *TubA1B* locus (Figure 2.13). Genotyping of these cells by PCR showed primarily heterozygous insertion of both mEGFP and mEos 3.2, with the exception being two mEos 3.2 clones (Clones 13 and 15) in Hel 92.1.7 cells.

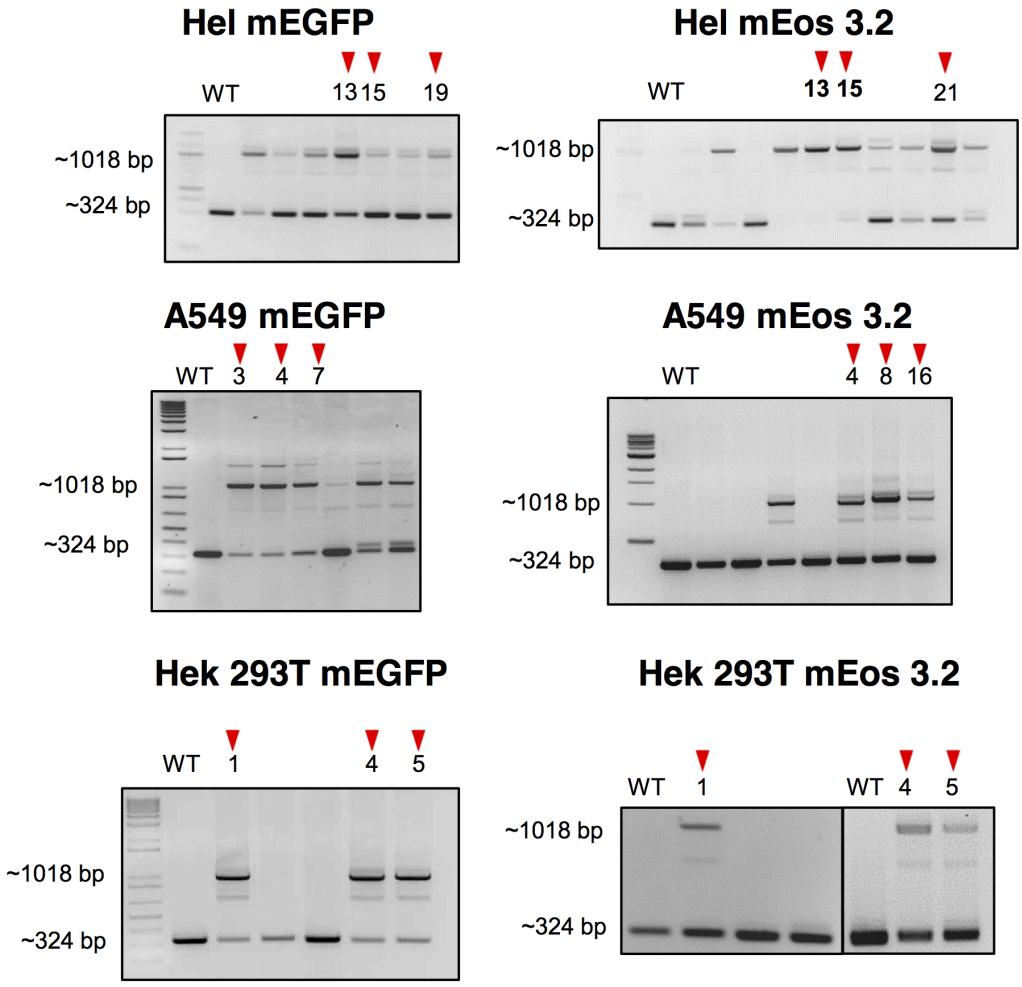

Figure 2.13: **Genotyping of single cell sorted *TubA1B* knock-in clones shows successful insertion of both fluorophores.** After single cell sorting and clonal expansion, DNA was extracted for mEGFP and mEos 3.2 clones for each of the three target cell lines (Hel 92.1.7, Hek293T, A549). As with the Dronpa clones previously reported, successfully targeted clones will feature a band at approximately 1000bp. Most clones are heterozygous for insertion, with the exception of two Hel 92.1.7 mEos 3.2 clones which demonstrate a homozygous knock-in and as a result the loss of a wild type band (Clones 13, 15).

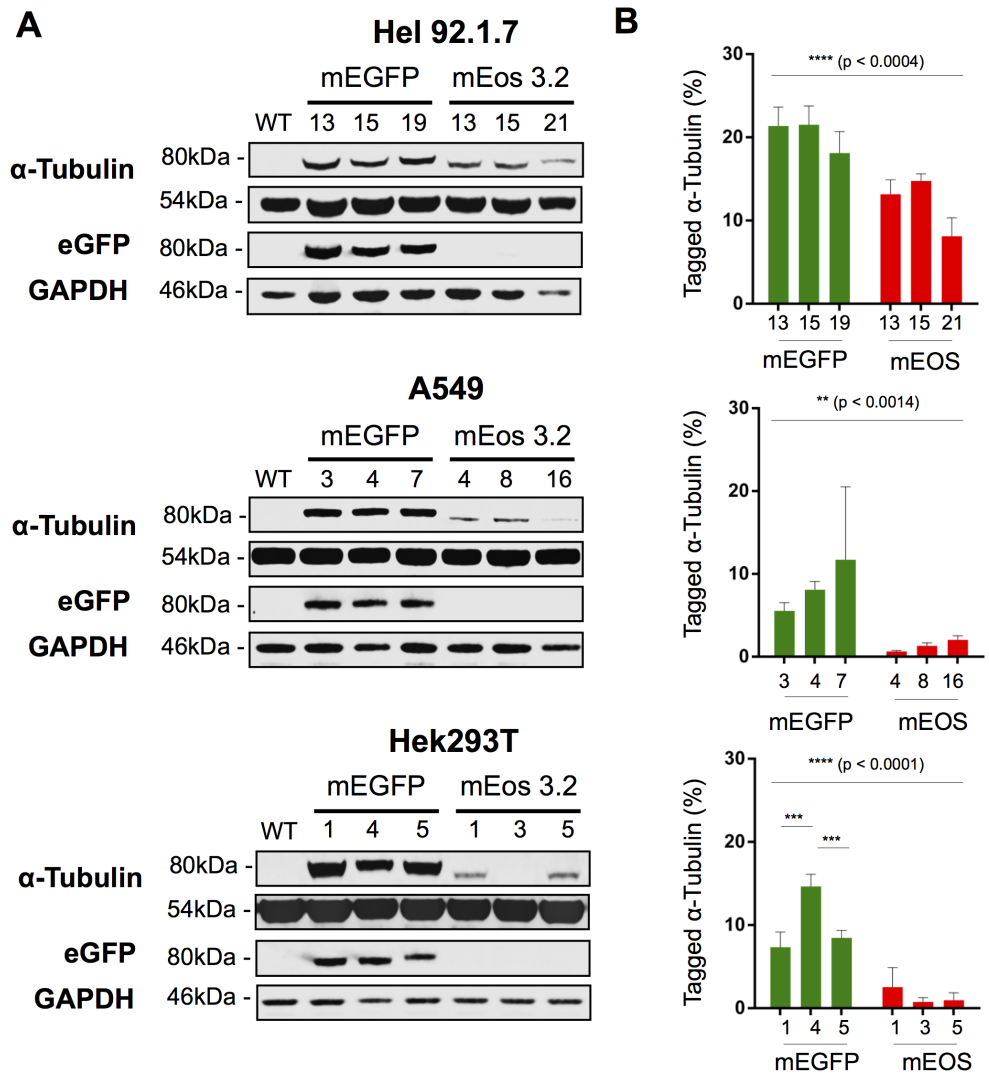

Figure 2.14: **Western blot analysis of *TubA1B* knock-in clones reveals a tag and cell line dependent expression of tagged tubulin.** PCR validated clones were expanded and lysates prepared for western blotting. (A) Western blot analysis of Hel 92.1.7, A549, and Hek293T clones tagged with mEGFP and mEos 3.2 reveals a successful genomic expression of the tagged protein, however with an obvious variation in tag expression across all three cell lines. (B) Quantification of these gels reveals both a tag and cell line dependent level of tagged *TubA1B* expression. Hel 92.1.7 consistently expresses the highest level of tagged α-tubulin, however there is a statistically significant difference in expression between tags (**** $p < 0.0004$). A549 and Hek293T cells show a markedly lower expression of tagged *TubA1B*, with mEos tagged cells from both lines showing a significant down regulation when compared to mEGFP equivalents (** $p < 0.0014$ and *** $p < 0.001$) respectively. $n = 3$, S.D., Two-Way ANOVA with multiple comparisons.

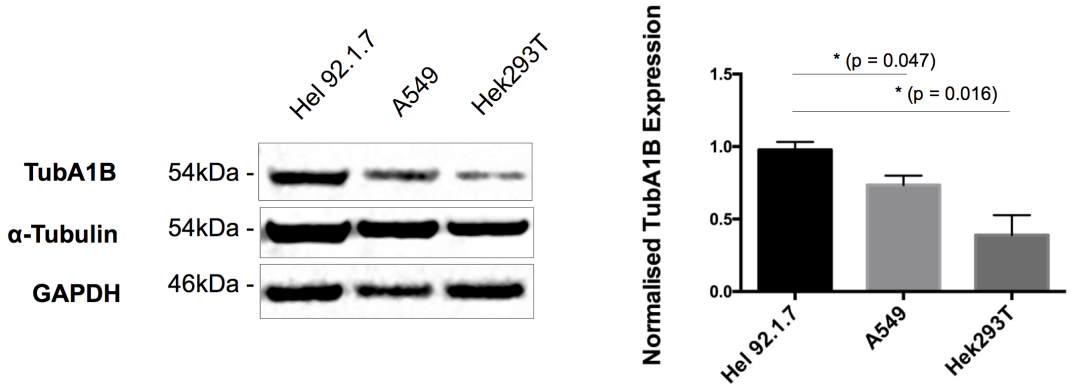

Figure 2.15: **Cell line specific expression of *TubA1B* impacts label distribution for SMLM imaging through endgenous expression.** Western blotting of wild type Hel 92.1.7, A549, and Hek293T lysate shows with a *TubA1B* specific antibody shows that Hel 92.1.7 cells express *TubA1B* higher than A549 and Hek293T.

Positive cells were taken forward for western blot analysis to quantify the expression level of tagged *TubA1B*. Interestingly, both a cell line and tag-dependent regulation of expression was observed (Figure 2.14A). Across all three cell lines, mEGFP tagged tubulin is well expressed, conversely mEos 3.2 tagged *TubA1B* is consistently down-regulated. A cell line dependent relationship is also observed, which is likely attributable to varying expression of this specific α-tubulin isoform in the cell lines studied (Figure 2.15).

Out of the three cell lines tested here, Hel 92.1.7 tagged clones most highly express tagged α-tubulin at approximately 20% for mEGFP and 15% in mEos 3.2 labelled clones. A549 cells express mEGFP tagged α tubulin at approximately 10% compared to 3% mEos 3.2, and finally Hek293T express mEGFP-*TubA1B* at between 8 and 15 compared to 2% for mEos 3.2. These findings are surprising as no previous work has demonstrated a sequence dependent downregulation of endogenously expressed fusion proteins.

Expression level is likely to have a significant effect on SMLM due to the stringent labelling requirements inherent to nanoscopic resolution. As demonstrated in figure 2.16, *TubA1B*-mEGFP folds correctly and assembles into distinct microtubules which co-stain with α-tubulin antibody. These cells are potentially powerful tools for long term live cell imaging using lightsheet modalities, where the high SNR can compensate for the relatively low brightness compared to over-expressed cells.

As expected however, the variation in expression has a substantial impact on the quality of CRISPR-PALM images as demonstrated in figure 2.17. Here CRISPR-PALM images are compared to dSTORM, and while Hel 92.1.7 cells demonstrate clearly resolved microtubules which compare favourably with dSTORM, the other two cell lines studied show punctate, incomplete structures when compared to the antibody based technique.

SMLM images can be compared through quantitative means to determine the quality of the data sets acquired. When comparing mean and mode localisation precisions, CRISPR-PALM in Hel 92.1.7 cells does not significantly differ from dSTORM data, however both A549

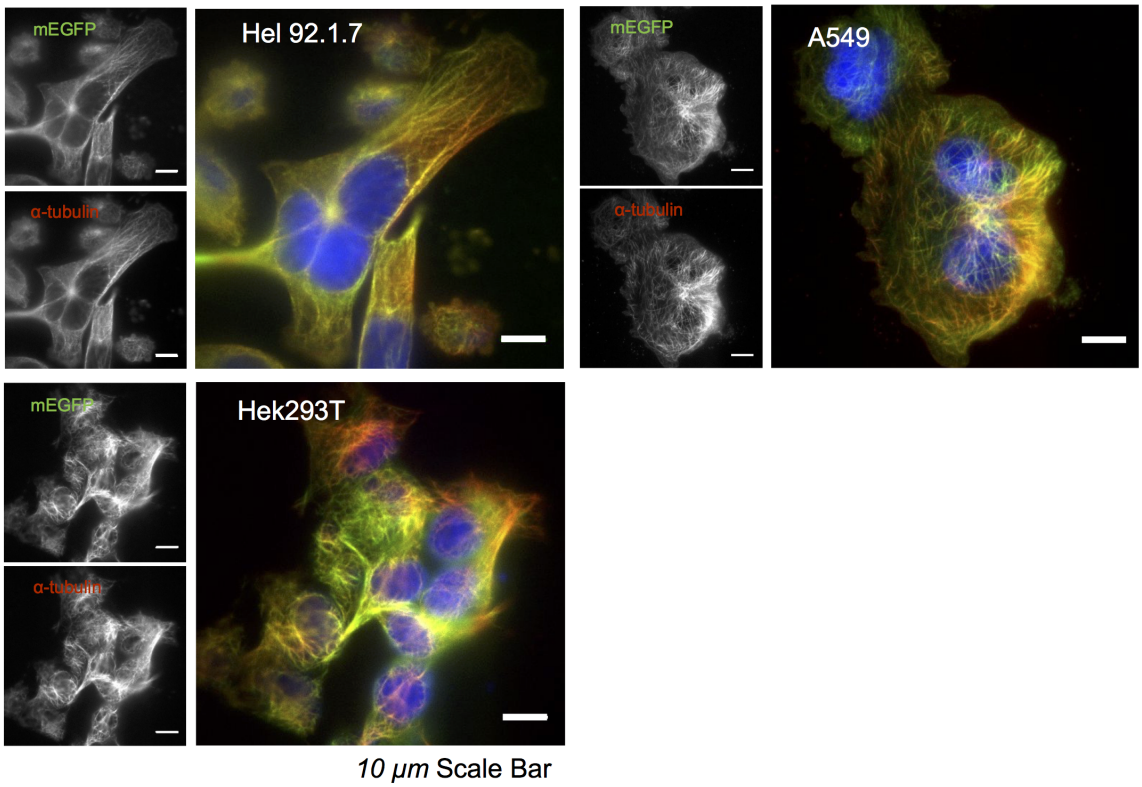

Figure 2.16: ***TubA1B*-mEGFP knock-in cell lines express correctly folded tubulin at sufficient brightness for diffraction limited imaging.** mEGFP *TubA1B* knock-in cells all show correct folding and expression of α-tubulin which co-stains well with an α-tubulin antibody. Despite the lower expression of the tagged protein in A549 and Hek293T, there is sufficient label density for diffraction limited imaging.

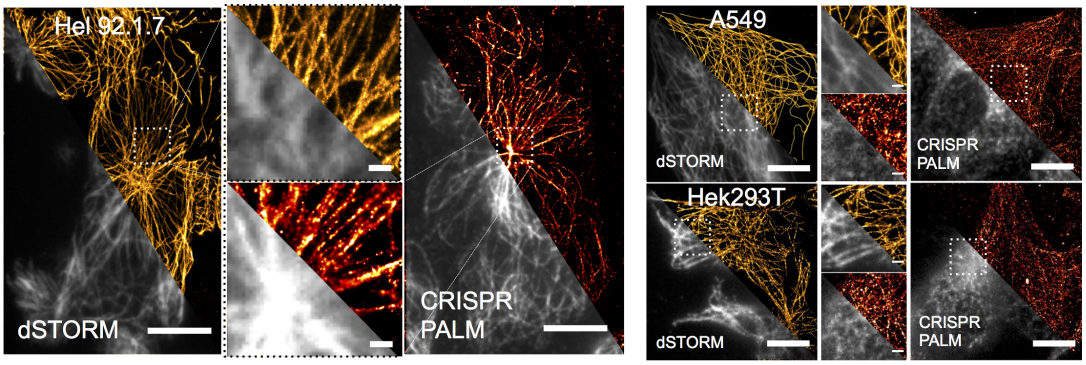

Figure 2.17: **Comparison of mEos 3.2 knock-in cells and dSTORM α-tubulin shows SMLM imaging dependent on cell line specific expression.** CRISPR-PALM (mEos 3.2 knock-in) imaging was performed in A549, Hek293T, and Hel 92.1.7 cells and compared to dSTORM images of α-tubulin labelled wild type cells. In A549 and Hek293T dSTORM imaging shows clearly resolved microtubules, while the respective CRISPR-PALM image suffers from a lack of sufficient labelling evidenced by poor, incomplete MTs. Interestingly, Hel 92.1.7 cells express enough mEos 3.2 labelled α-tubulin to achieve comparable SMLM images.

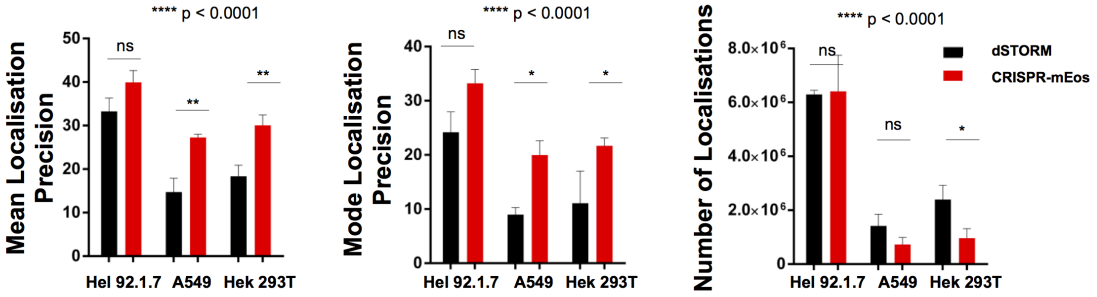

Figure 2.18: **Quantitative comparison of CRISPR-PALM and dSTORM confirms cell line specific quality of SMLM imaging.** A quantitative comparison of CRISPR-PALM through knock-in of mEos 3.2 with dSTORM images shows that the expression level has a substantial effect on the quality of subsequent single molecule imaging. Two key metrics are compared here, localisation precision (mean and mode) and total number of localisations. Across all three measurements, a significant variation is observed across cell lines (**** p ¡ 0.0001). In Hel 92.1.7 cells, which highly express the mEos 3.2 tagged tubulin, mean and mode localisation precision and total number of localisations are not significantly different when comparison to dSTORM. In A549 and Hek293T however, a statistically significant difference in all measurements is evident. This is consistent with poor quality SMLM as a result of the down regulation of the mEos 3.2 tagged allele. $n = 3$, S.D. Statistical testing performed using a 2-Way ANOVA with multiple comparisons.

and Hek293T cells are significantly poorer compared to dSTORM (**, * for mean and modal precision respectively according to multiple comparisons). While the higher photon budget of organic fluorophores does naturally improve the localisation precision of subsequent fits, the precision measured in ThunderSTORM reconstructions is also affected by background and out of focus detections which are likely to be significant in antibody labelled samples. Similarly the total number of localisations is significantly reduced in CRISPR-PALM A549 and Hek293T cells when compared to dSTORM, while there is no significant difference between the number of localisations in Hel 92.1.7 samples.

Finally measures of effective resolution can be applied to compare the quality of CRISPR-PALM to dSTORM. A collection of full width half maxima (FWHM) of microtubules across all cells observed shows no significant difference in FWHM between dSTORM and CRISPR-PALM in Hel 92.1.7 cells (Figure 2.19) ($p = 0.086$). Using a relatively recent measure of effective resolution called the Fourier Ring Correlation method a slight improvement in Fourier Image Resolution (FIRE) can be observed in CRISPR-PALM cells ($p = 0.0447$)[215]. This is most likely due to the more even labelling density observed in CRISPR-PALM cells (as FIRE calculation takes sampling density into account).

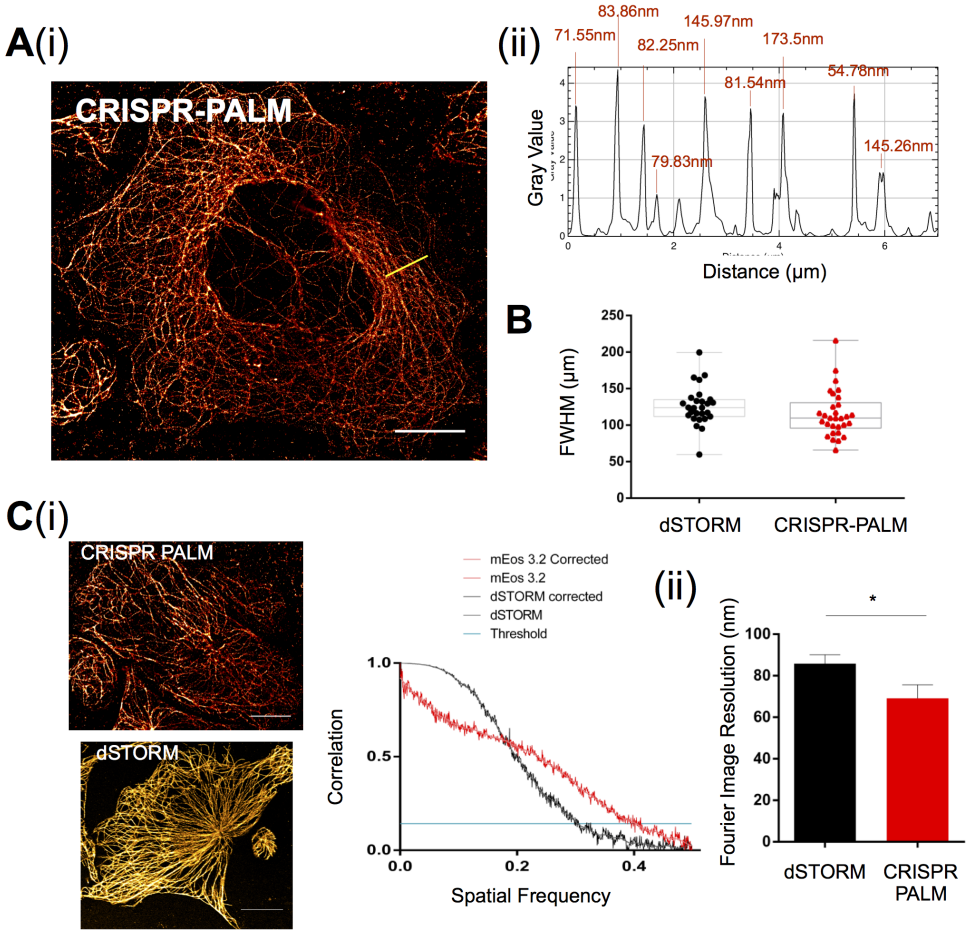

Figure 2.19: **Effective resolution in CRISPR tagged Hel 92.1.7 *TubA1B* is comparable to dSTORM.** CRISPR-PALM images of Hel 92.1.7 were further evaluated to determine how effective resolution of this modality compares to dSTORM. (A) (i) a representative line profile shows (ii) clearly resolved peaks. When these are compared to dSTORM images across multiple replicates, no significant difference is observed between the two modalities. ($p = 0.086$) (C) A comparison of Fourier ring correlation (FRC) as a measure of effective resolution shows a significant improvement in CRISPR-PALM when compared to dSTORM, presumably due to the consistent labelling density in the genomic expression of tagged protein ($p = 0.0447$). $n = 3$, S.D. Statistical testing performed using unpaired t-test.

2.3.4 Applications of CRISPR-PALM cells

Multiplexing has been a significant issue in the field of SMLM due to the substantial difference in fluorophore performance - both in PALM (between photoswitchable fluorophores) and dSTORM (between organic fluorophores). A CRISPR-PALM approach offers a relatively simple method of combining high performing dSTORM fluorophores (e.g. Alexa647) with well characterised PALM tags (mEos 3.2), without the loss of performance in particular blinking buffers. In figure 2.20 Alexa-647 labelled TTLL10, a tubulin ligase, is imaged in a Hel 92.1.7 clone edited to express *TubA1B*-mEos 3.2. The result is high quality SMLM where spatial relationships between tubulin and its accessory protein are revealed at nanoscale resolutions.

Another significant advantage of CRISPR-PALM is the option of live cell imaging, particularly over periods of time which would exceed those allowed by transient transfection (48-72 hours). Significant concerns in live localisation microscopy include phototoxicity and temporal resolution - while classical PALM would require millions of localisation events to be recorded within each acquisition, more recent advances which utilise non-linearity in photoswitching data sets can super-resolve without over exposure to light (and hence toxicity), or a loss of temporal resolution. Most significant of these is the recent development of SRRF (Sub-pixel radial fluctuations), an approach which uses sub-pixel radial symmetry to super-resolve at sub-second temporal resolutions and low light intensities (Figure 2.21).

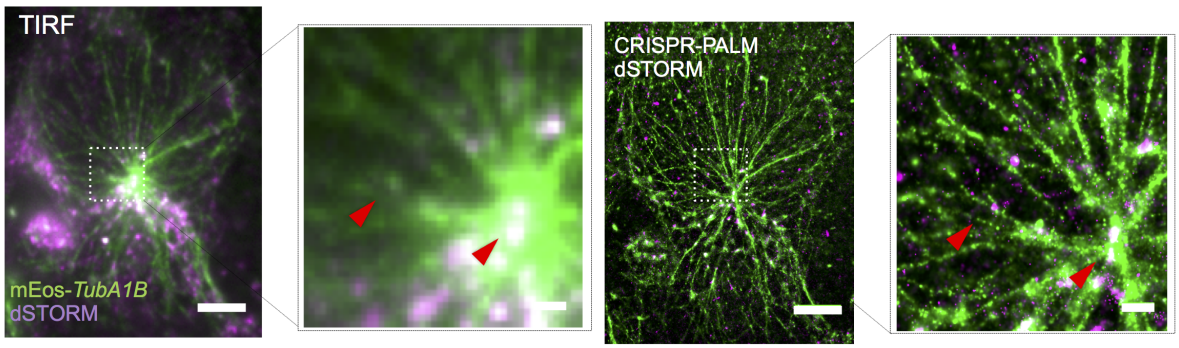

Figure 2.20: **Multiplexing CRISPR-PALM and dSTORM for convenient two-colour SMLM.** Multiplexing SMLM tags has classically been a challenge due to the diverse photochemistry and blinking kinetics of organic fluorophores. The use of a CRISPR-PALM approach compliments the use of high performance dSTORM antibodies for high quality two-colour imaging without the need to compromise on one of the two reporters used. Here two-colour imaging is used to achieve nanoscopic resolution of both *TubA1B* (green) and TTLL10, a tubulin ligase (magenta).

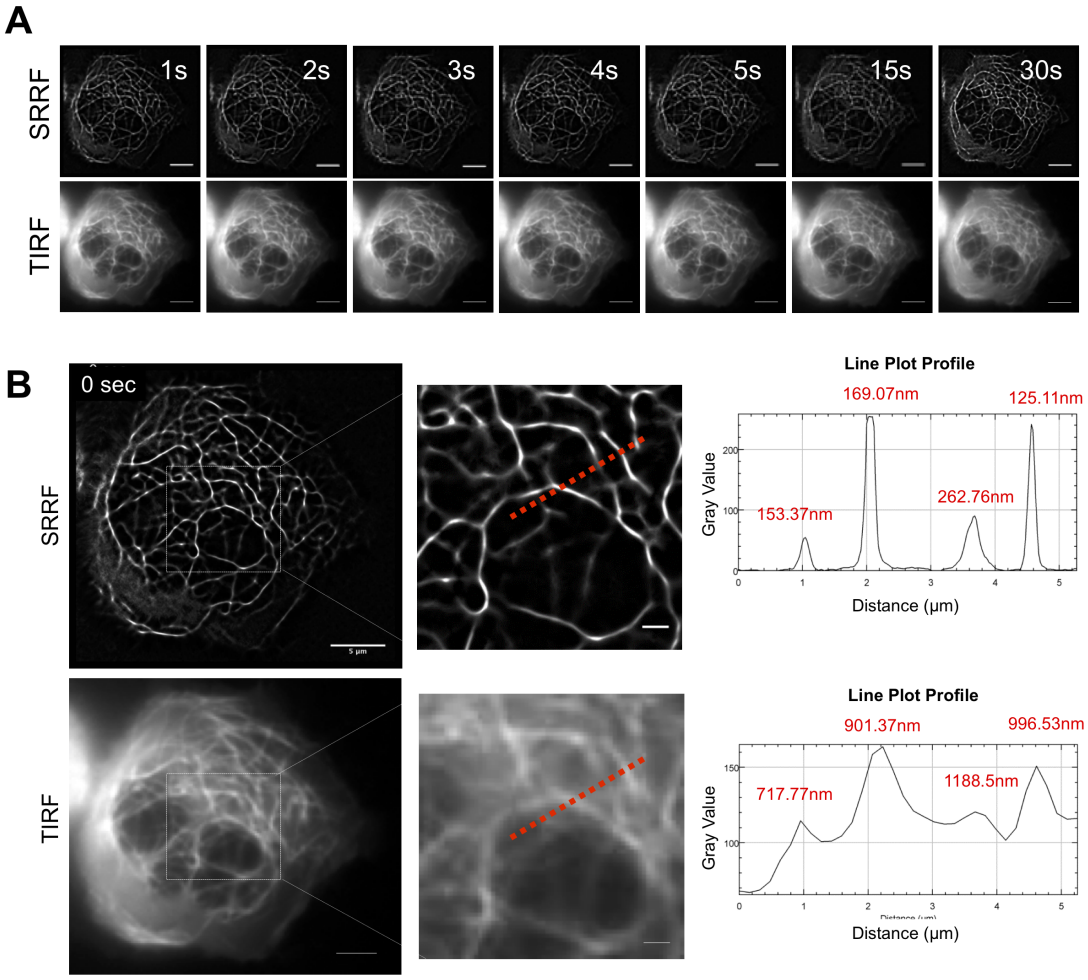

Figure 2.21: **CRISPR tagged mEos 3.2-*TubA1B* Hel 92.1.7 cells allow for super-resolved live cell imaging.** An endogenous expression system offers many other benefits, including the capacity for live cell imaging. (A) A recent publication applied the use of sub-pixel fluctuations to super-resolve at low laser intensities. This approach, called SRRF, can be used on mEos 3.2 knock-in Hel 92.1.7 cells to image the endogenously labelled *TubA1B* gene at sub-second temporal resolutions. (B) FWHM from a sample frame show effective super-resolution in the SRRF image compared to an equivalent averaged widefield frame.

2.3.5 Fluorophore properties directly impact protein expression after CRISPR-Knock-in

The sequence specific downregulation of tagged *TubA1B* observed in our knock-in cells is a surprising finding with significant consequences for the field of CRISPR mediated knock-in. To determine whether fundamental properties of the fluorophore and its sequence drive this effect, three mEos variants were designed and cloned into the donor plasmid (Figure 2.22A). A codon optimised mEos 3.2 variant, as well as the recently reported mEos 4b and a codon optimised variant of the original sequence. mEos 4b has been reported as a truly monomeric variant of mEos[216]. Each of these fluorophore sequences were cloned as previously described using Gibson assembly, and after clonal isolation and expansion, digested with EcoRI to validate the presence of an insert (Figure 2.22B). Sequence alignments as shown in figure 2.22 indicate variations in the 4b sequence from the original 3.2 insert, and that codon optimisation does not alter the protein sequence of these inserts.

Sequence validated donors were then co-transfected in Hel 92.1.7 cells with a px459 *TubA1B*-guide C plasmid and single cell sorting (Figure 2.23A). As Hel 92.1.7 cells best express *TubA1B*, they provide a good sample in which to try and reverse the downregulation of mEos observed. As shown in figure 2.23A, wild type cells are negative for signal in the FITC channel, while validated mEos 3.2 expressing clone 13 is positive for signal in this channel. Blots for mEos 4b and its codon optimised variant show a higher rate of insertion compared to the codon optimised mEos 3.2 (5.31% and 5.70% respectively compared to 1.9%).

After single cell isolation and expansion, individual clones were genotyped and validated by PCR amplifying the targeted region (Figure 2.23B). Once again positive clones are predominantly heterozygous (with one exception - mEos 4b clone F8), demonstrating a band approximately 700bp heavier than wild type as expected.

Validated clones were then subjected to western blot analysis to determine whether

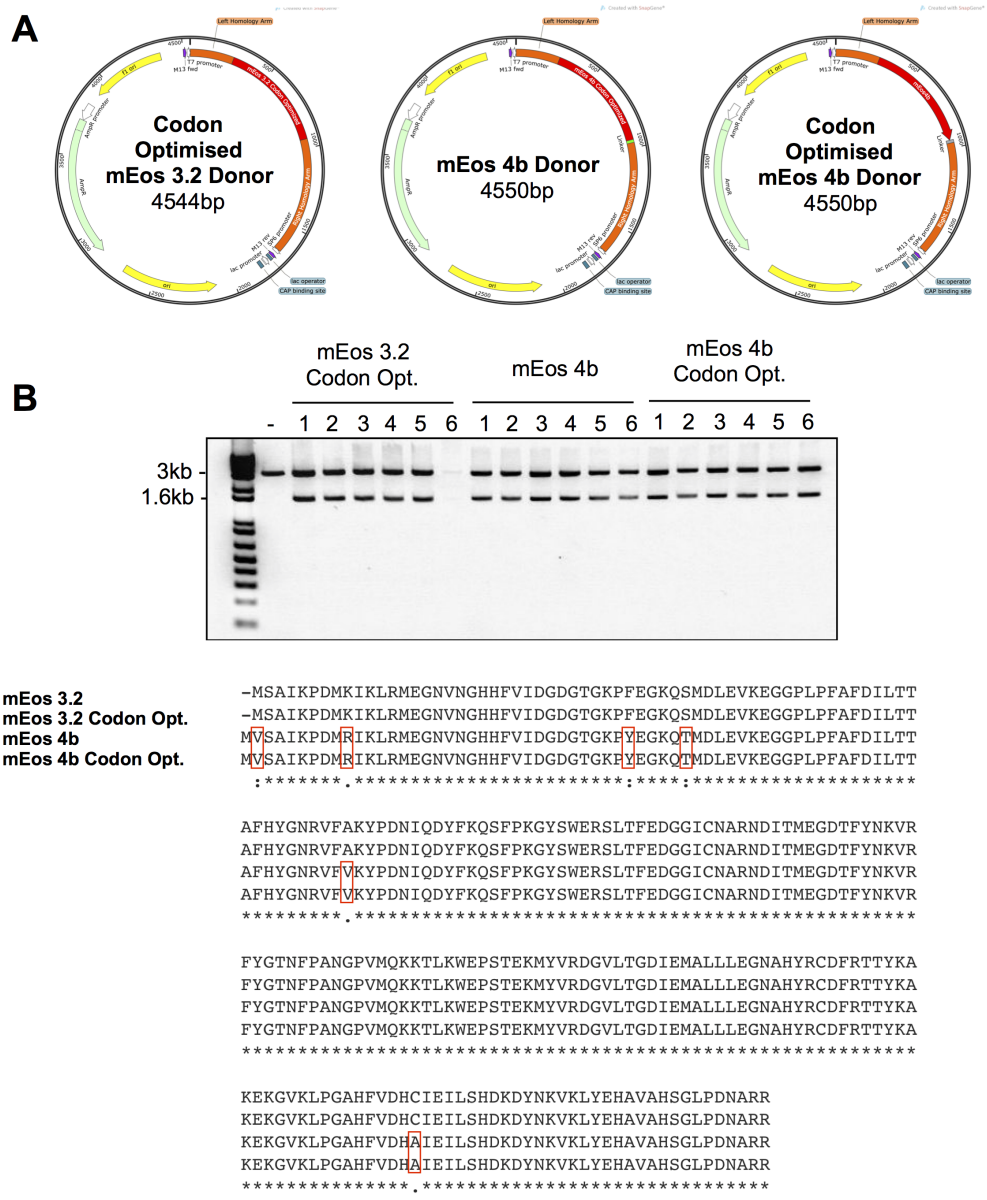

Figure 2.22: **Generation of *TubA1B* donor plasmids with mEos variants.** (A) Donor plasmids with identical linkers and homology arms harbouring mEos variants were designed. These include a codon optimised mEos 3.2 sequence, as well as a truly monomeric version of the protein reported as mEos 4b, and a respective codon optimised variant. (B) gBlocks were cloned into pGem-T-Easy backbones and validated through EcoRI digest. Negative controls are 3000bp in length, while successfully cloned vectors demonstrate an insertion which runs at approximagely 1600bp. Finally sequence alignments show differences in the 3.2 and 4b protein sequence, and that codon optimisation does not alter the reading frame or coding sequence reported. Amino acids marked in red are where the mEos 4b protein differs from its predecessor mEos 3.2.

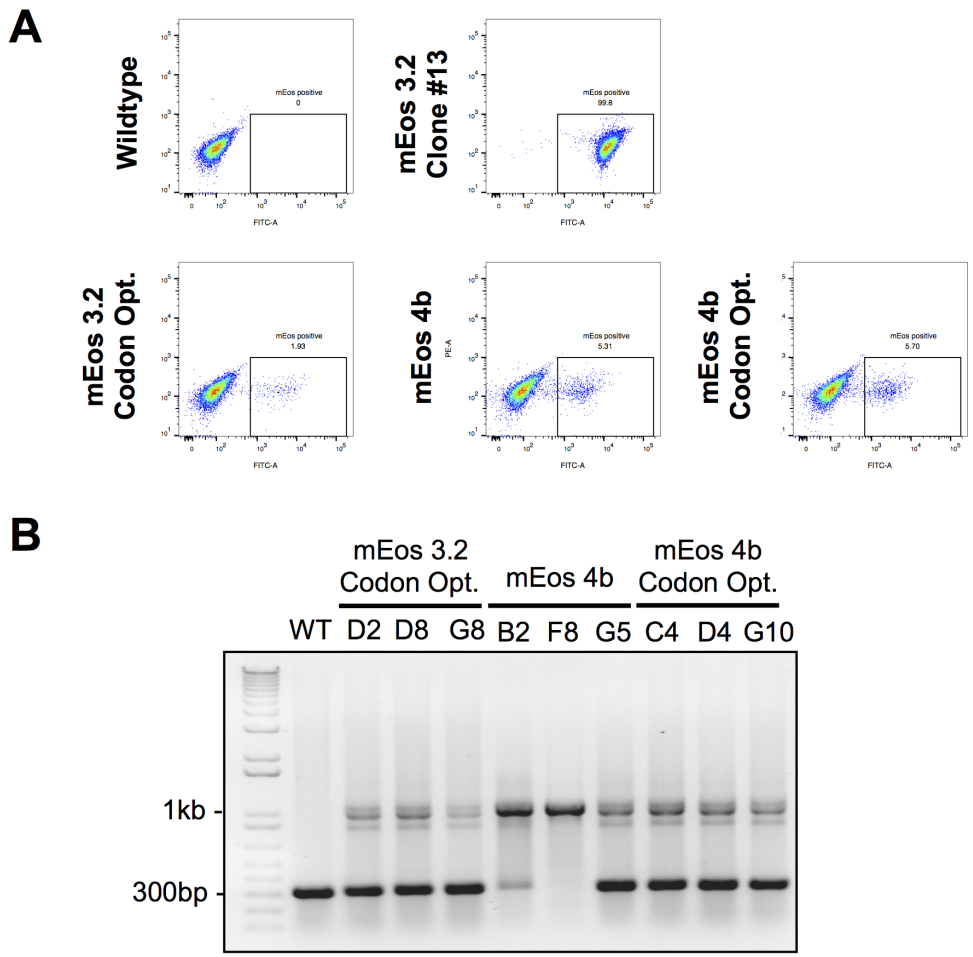

Figure 2.23: **Single cell sorting and PCR based genotyping of Hel 92.1.7 clones carrying mEos variants to determine successful integration.** (A) Hel 92.1.7 cells co-transfected with a guide targeting *TubA1B* and donor templates carrying codon optimised mEos 3.2, mEos 4b, and codon optimised mEos 4b were single cell sorted into a 96 well plate. Cells were compared to previously generated mEos 3.2 clone 13 which stably expresses *TubA1B*-mEos 3.2 fusion as a knock-in, and is therefore a completely fluorescent population and a useful positive control. (B) Clones were expanded and validated by PCR. The amplification of wild type DNA results in a 300bp band, while heterozygous clones positive for fluorophore insertion at the *TubA1B* locus also feature a second band at approximately 1000bp.

varying the properties of mEos and its sequence impact the expression of tagged α-tubulin (Figure 2.24A). Quantification of these western blots shows that while there is no significant difference between codon optimised mEos 3.2 and the parent sequence, there are significant differences between the mEos 3.2 clone and both mEos 4b and mEos 4b codon optimised. Surprisingly, expression of mEos 4b is significantly reduced across all three clones (*** $p = 0.0016$, $p = 0.0015$, $p = 0.0023$ for each clone respectively), but elevated in 2 of the three codon optimised mEos 4b knock-in cells (** $p = 0.0084$, ns $p = 0.1$, * $p = 0.011$ respectively).

The improved expression of the codon optimised mEos 4b knock-in suggests that a highly monomeric, codon optimised of mEos is better suited to SMLM due to the improved abundance and distribution of labelled protein. PALM performed on codon optimised mEos 4b clone C4 results in high quality super-resolution imaging (Figure 2.24C), with significantly more localisations observed in total and across each frame when compared to mEos 3.2 knock-ins (* $p = 0.0230$) (Figure 2.24C,D). There is no significant difference in FRC or XY uncertainty, consistent with the fact that despite the improved expression, mEos variants still demonstrate the same fundamental fluorophore properties (molecular brightness, quantum yield, switching kinetics etc.) ($p = 0.3844$ and $p = 0.0861$ respectively) (Figure 2.24E, F).

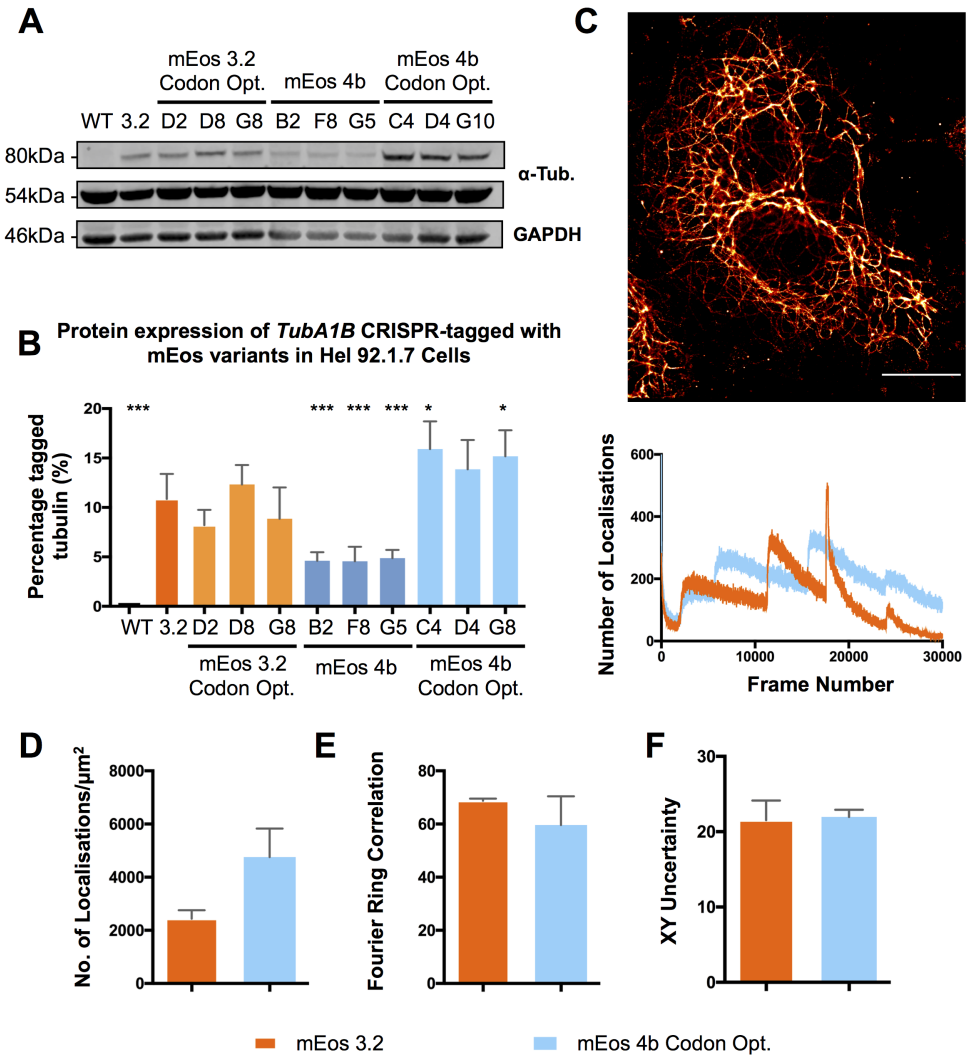

Figure 2.24: **Improved expression of a codon optimised, monomeric mEos variant results in high quality CRISPR-PALM.** Hel 92.1.7 cells expressing mEos variants are compared to mEos 3.2 knock-ins by western blot and imaging. (A) Western blotting shows a substantial improvement in tagged tubulin expression in cells labelled with codon optimised mEos 4b when compared to other mEos variants. (B) Quantification of western blotting shows no significant difference in expression when comparing codon optimised mEos 3.2 and the original mEos 3.2 knock-in. mEos 4b demonstrates a significant reduction in expression (** $p = 0.0016$, $p = 0.0015$, $p = 0.0023$ for each clone respectively), while codon optimised mEos 4b shows significantly improved expression when compared to mEos 3.2 (** $p = 0.0084$, ns $p = 0.1$, * $p = 0.011$ respectively). (C) The improved expression of mEos 4b CO results in a higher number of detections reported per frame compared to mEos 3.2. (D) Codon optimised mEos 4b demonstrates a significantly higher number of localisations when compared to mEos 3.2 knock-in clones (* $p = 0.0230$). (E) No significant difference in fourier ring correlation is observed ($p = 0.3844$). (F) Similarly no difference in xy uncertainty is observed as expected from mEos variants ($p = 0.0861$). ($n = 3$. S.D., Two-Way ANOVA with multiple comparisons for western blot experiments. Unpaired t-tests performed for uncertainty and number of localisations.)

2.3.6 Towards a two-colour CRISPR-PALM approach

Finally to begin developing CRISPR-PALM as a two colour approach, *TubA1B* donors carrying a HaloTag™insert were generated. HaloTag™refers to a modified bacterial enzyme (a haloalkane dehalogenase) is capable of hydrolysing choloroalkane linkers on synthetic ligands or organic fluorophores [217, 218]. The enzyme can be expressed instead of a fluorescent protein, and is therefore not fluorescent in its native state. However, on incubation with a label conjugated to an appropriate chloroalkane linker, the enzyme hydrolyses the linker to effectively bind the reporter.

This approach allows for genomic expression, and thus the control of labelling stoichiometry, as well as a a range of different choices in terms of fluorophores and the advantageous photon counts typical of organic fluorophores[217, 218]. Cell impermeable labels (e.g. Alexa Fluor, Atto-488) can be used to label membrane bound elements and image their internalisation, while permeable fluorophores (e.g. TMRs, Janelia Fluor) can be used to label intracellular components. Such an approach offers a great deal of experimental flexibility and the benefits of both expression mediated labelling and the brightness and photostability of organic fluorophores[217].

As previously described, a codon optimised HaloTag™sequence was cloned into a *TubA1B* donor plasmid and validated (Figure 2.25A,B) before co-transfection into Hel 92.1.7 cells with *TubA1B* targeting guide C. HaloTag™cells were single cell sorted after incubation with Janelia Fluor 546 (which appears positive in the PE channel) (Figure 2.25C).

Single cell clones were imaged with JF546 and evaluated for expression using western blot. While the cells demonstrate tagged α-tubulin comparable to mEGFP (Figure 2.26A), western blot based quantification of tagged protein shows low levels of expression (approximately 5%), substantially lower than both mEGFP and any mEos variant.

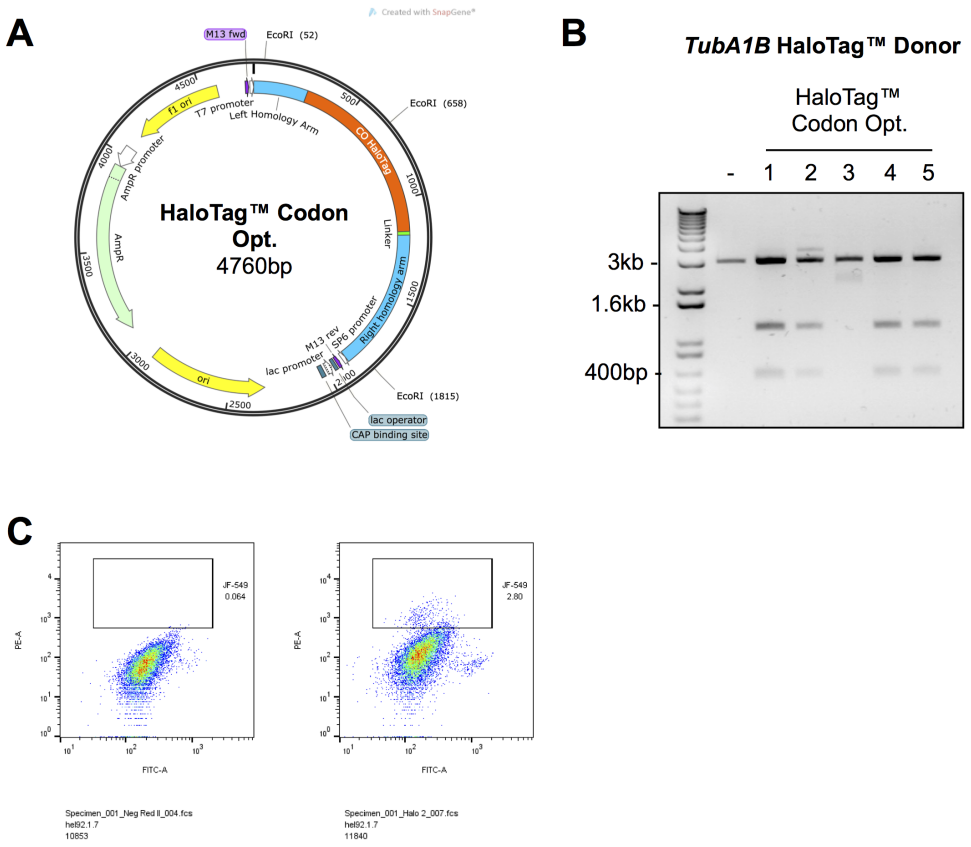

Figure 2.25: **Generation of donor plasmid carrying codon optimised HaloTag™and subsequent single cell sorting post-transfection.** (A) A donor vector carrying a codon optimised HaloTag insert flanked by *TubA1B* specific homology arms was designed and cloned into a pGem-T-Easy backbone. (B) Clones were validated by EcoRI digest before sequencing to confirm correct assembly. (C) Validated *TubA1B* HaloTag™donors were then co-transfected with a guide and Cas9 expressing plasmid before single cell sorting.

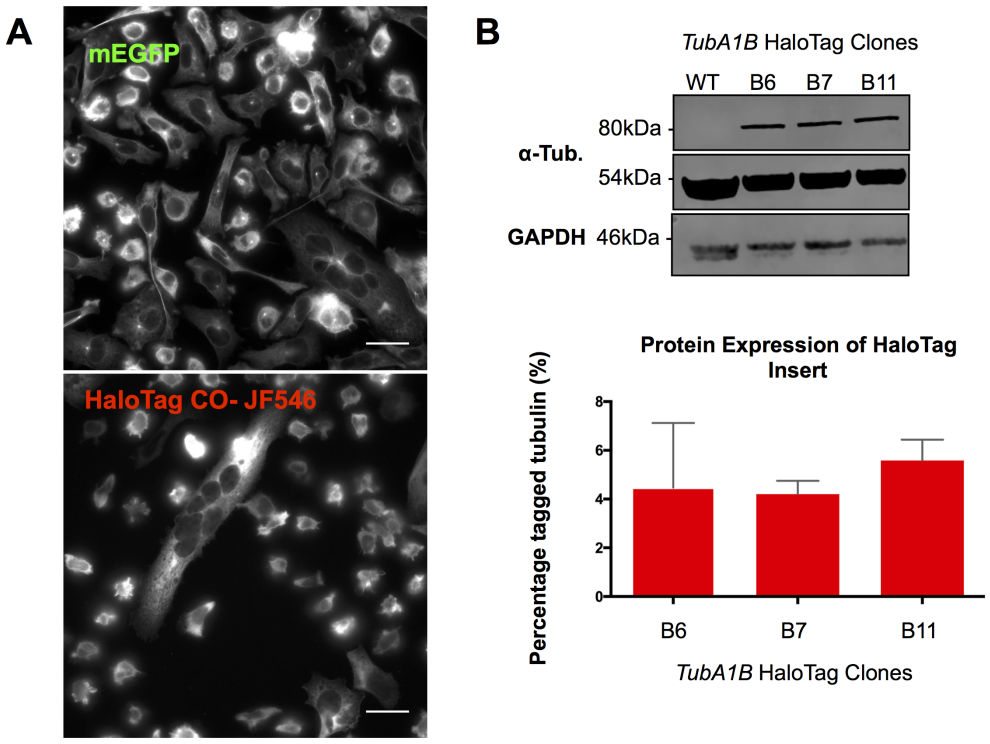

Figure 2.26: **Knock-in of HaloTag™ at the *TubA1B* locus in Hel 92.1.7 cells results in a substantial down-regulation of the tagged allele.** (A) *TubA1B* HaloTag™ knock-in clones fold correctly and are comparable to mEGFP clones after incubation with Janelia Fluor 646. (B) Expression of HaloTag™ was evaluated by western blot and found to be low (approximately 5%) compared to previously reported expression levels of other tags.

2.4 Discussion

The use of CRISPR-PALM as a means by which to generate SMLM data sets has the potential to address a number of long standing issues in the field of single molecule imaging. Most significantly perhaps, a known stoichiometry promises accurate quantification, particularly when combined with the known switching kinetics of a photoswitchable tag. The use of endogenous expression would overcome the issues associated with transfection inherent to PALM, and allow for super-resolution imaging of novel genes for which quality antibodies would likely be unavailable.

However the caveat to endogenous expression will be the loss of labelled protein in the sample - previous attempts at PALM have always benefited from saturating cells with labelled protein to satisfy the strict labelling/sampling criteria of single molecule localisation microscopy. A challenge in developing a single molecule system based on endogenous expressed will be expressing enough labelled protein to achieve nanoscopic resolutions - in the work reported thus far, this is further complicated by an apparent sequence dependent down regulation of tagged tubulin.

This work presents an efficient method of cloning donor sequences to generate CRISPR tagged knock-in lines. The use of gBlocks (synthetic DNA) and Gibson assembly has allowed for the generation of multiple donor sequences with relative ease. A host of other donor plasmids carrying alternative tags not presented here have been produced for future work - including rsKame, rsFastLime, mApple, PAmCherry, and rsTagRFP. Future work will involve comparing knock-ins of these to determine which fluorophores are best expressed and what properties determine this.

The work presented in this chapter demonstrates the first attempt to establish a CRISPR based method of single molecule imaging. Microtubules are an ideal test subject as they present a number of challenges which, if overcome, can yield high quality images. As structures well beneath the diffraction limit (25nm diameter) they are ideally suited for

SMLM, however to achieve resolution in the tens of nanometers sufficient labelling of these cytoskeletal structures needs to be achieved.

The lack of homozoygous knock-ins across all three cell lines has been surprising - we reason that this is likely to affect microtubule formation, and hence cell viability. At the time of the publication of this work, no discussion of the potentially deleterious effects of bi-allelic insertion was found in the literature. Since then, work by the Allen Institute in the generation of stem cell libraries independently confirmed that it was not possible to produce bi-allelic *TubA1B* CRISPR knock-ins[219]. Roberts *et al.* report a generally low incidence of bi-allelic insertions across a number of genes. Interestingly in one of the only other approaches designed to apply knock-ins for super resolution microscopy (RESOLFT), Ratz *et al.* report bi-allelic insertions for two of their genes (ZYX and HMGA), but only mono-allelic knock-in for their cytoskeletal target, Vimentin. While the authors do not quantify their knock-in of rsEGFP2, their data suggests that they too are restricted to mono-allelic insertions for this target.

This finding has important implications for the application of CRISPR-Cas9 for HDR mediated fluorophore insertion, particularly in the context of the sequence and fluorophore dependent down regulation of tagged alleles reported in this chapter. To our knowledge, no work to date has reported this phenomenon - however it stands to reason that there are cellular mechanisms which maintain morphology and function at the cost of the translation of a potentially detrimental gene. In this work, it would appear that monomericity (mEGFP and mEos 4b are highly monomeric) and codon optimisation (mEGFP and codon optimised mEos 4b) are important factors in determining the expression of tagged alleles[220, 216]. Interestingly, codon optimisation of mEos 3.2 is insufficient to achieve high levels of expression, suggesting that in this instance the monomericity of the protein restricts expression.

Comparisons between transfected and knock-in cells show that knock-ins maintain a cells native morphology and function[221, 206]. In work by Uemura *et al.*, endogenous expression of

an EGFP β-actin fusion does not alter dendritic spine morphology or function - in contrast, over-expressing the equivalent fusion protein results in aberrant morphology and growth. Similarly Ratz *et al.* demonstrate an alteration of morphology and protein distribution in over-expression cells when compared to equivalent knock-ins. In the context of the work presented in this chapter, it is likely that knock-in cell lines down-regulate protein to maintain morphology and function. To our knowledge, this has not been explicitly reported in the literature. Similarly the sequence dependent effect observed here has not, to our knowledge, been reported, with significant implications for the wider application of CRISPR knock-in.

On one hand, the preservation of function at the cost of expression is biologically advantageous. A long standing criticism of transfection has been the aberrant expression of fusion proteins, and hence, biological artefacts as a result. In quantitative SMLM, which has particular applications in receptor biology, this can lead to clustering artefacts dependent on expression level. CRISPR knock-in offers a means by which to ameliorate this by allowing cells to regulate the genomic expression of fusion proteins in the interest of biological integrity. Unfortunately, a caveat to this will be the loss of signal which is critical to high quality live cell imaging. It is also likely that some genes will not tolerate genomically expressed fusion proteins, and that some fluorescent proteins will have a detrimental effect on expression as we have observed in this work.

While we have observed that high quality SMLM comparable to dSTORM can be achieved through CRISPR-PALM, the sequence dependent down-regulation of tagged alleles is an extremely important observation. While this work establishes a role for monomericity and codon optimisation in altering endogenous expression, a host of other important fluorophore parameters have yet to be explored. It is likely that maturation rates, size, and linker design may play a role in regulating expression level. Future work in this and other labs should aim to develop fluorophores optimised for CRISPR knock-in. Similarly the careful quantification of protein expression after knock-in should become best practice after CRISPR knock-in.

Chapter 3

Development of iPSC Protocols

The lack of platelet producing cell lines has been a significant obstacle in the field of megakaryocyte and platelet biology. The absence of a reliable, high throughput *in vitro* approach has restricted much of the work in MK and platelet biology to mouse models, patient samples, and relatively poor leukaemic cell lines. In the context of the study of inherited thrombocytopenias, and in particular the GAPP study, this is a significant obstacle when studying novel genes and/or mutations.

In this chapter two key technologies are discussed in the context of their applications in this field, namely induced pluripotent stem cell (iPSC) differentiation and Clustered Repeating Interspaced Palindromic Repeat (CRISPR) gene editing. After a brief introduction, efforts to generate differentiation models and to introduce CRISPR knock-ins to stem cells are reported.

3.1 Chapter Introduction

While a number of cell lines exist which express a range of MK markers, no existing immortalised cell line produces proplatelet extensions or platelets *in vitro*[222]. As such murine bone marrow and foetal liver derived megakaryocytes have been a staple tool in interrogating

platelet production[223, 224, 110]. That being said, the generation of mouse models is an expensive and time consuming business which as such, is by its nature, low throughput.

Human cord and peripheral blood derived megakaryocytes have been one of the main methods of obtaining functional, human megakaryocytes to study. Once isolated from CD34+ cells, mature megakaryocytes can be generated after approximately 12 days of culture. These cells can then proceed to form proplatelet structures *in vitro* in a substrate and time dependent manner [225].

However the yield of proplatelet forming MKs has always been very low in these human derived cells, and while they can be virally manipulated to introduce reporter proteins and mutant phenotypes this process is often detrimental to cell viability[226]. In cultures where the number of cells derived is usually quite low, the effect of viral manipulation on sample size can have a significant effect on the quality and throughput of experiments.

With these limitations in mind, stem cell technologies offer a solution to the problem of a robust, high throughput *in vitro* assay to screen mutations and proteins of interest in mature, stem cell derived MKs. The key question is whether current methods of iPSC differentiation can yield the numbers of proplatelet producing cells needed for the quantitative phenotypic study of mutations of interest.

3.1.1 Methods of iPSC derived megakaryocyte differentiation

In one of the most significant works of modern regenerative medicine, Takahashi *et al.* published the successful reprogramming of adult mouse fibroblasts to form colonies with stem cell like pluripotency, later dubbed induced pluripotent stem cells (iPSCs)[227].

Plasticity in somatic cells had been previously shown, first in the 1960s where the injection of a nucleus from an adult somatic cell into an enucleated oocyte results in the development of a healthy frog[228, 229]. These findings were later corroborated by the successful cloning of 'Dolly' the sheep as a phenomenon reproducible in mammals[230]. These somatic cell

nuclear transfer (SCNT) approaches challenged the dogma of terminal differentiation, and demonstrated the presence of re-programming factors with the capacity to introduce the epigenetic changes needed for pluripotency. Takahashi *et al.*'s seminal work resulted in a significant leap forward as reprogrammed adult stem cells overcome many of the ethical and logistical issues inherent for the use of embryonic human stem cells[231].

In the field of megakaryocyte and platelet biology iPSC differentiation presents an opportunity to develop human, mature, proplatelet producing cells *in vitro*[232]. The driving force behind the development of mature iPSC MKs is the potential for the generation of large numbers of potentially transfusable platelets[232]. As anucleate cells in constant demand in blood banks globally, stem cell derived platelets are a viable option for the first *ex vivo* cell type generated for transfusion[233].

Directed differentiation is one of the key methods by which iPSCs can be driven towards a particular cell lineage. These approaches typically depend on the use of a substrate to support differentiation and the introduction of an array of cytokines designed to drive differentiation towards a germ layer. To generate MKs, this method is based on the induction of medodermal differentiation through BMP4 (Bone Morphogenic Protein-4) and Wnt signalling[234]. This mesoderm like layer is further driven towards haematopoeitic differentiation by the formation of haemogenic endothelium (or a haematoendothelial mesoderm). Early work applied this method to the production of MKs and platelets from human Embryonic Stem Cells and on feeder layers[235]. While effective, these approaches are limited in their applicability in transfusion medicine (where animal derived culture components are strictly forbidden and large scale stem cell culture and production are needed), and as such later work has focussed on applying these techniques to induced pluripotent stem cells[235, 236, 237].

Feng *et al.* applied directed differentiation to generate a large population of mature iPSC derived MKs in a feeder free format, however these cells generate an extremely small number of platelets (approximately 6 per MK) compared to the several thousand produced *in vivo*[238].

The authors drive pluripotent cells towards a mesoderm lineage by plating on Collagen IV and incubating cells in a phase I medium containing VEGF, FGF, and BMP-4 at 5% oxygen. After 6 days of culture, the authors incubate their haemogenic endothelium in a second cocktail of cytokines including factors critical for the expansion of haematopoeitic progenitors. The authors report a population which is approximately 60% positive for CD31 and 35% positive for CD34, both key markers of haemogenic endothelium.

The cytokine cocktail includes SCF, TPO, IL-3, IL-6, and Flt-3, and results in the generation of haematopoeitic progenitors (CD34+ cells), which the authors freeze and terminally differentiate to produce a population of CD41/CD42+ MKs (approximately 40%). These cells then go on to produce proplatelet extensions and platelet like particles which appear to respond to Thrombin (albeit an impaired response compared to *ex vivo* human platelets even at high doses (1U)), spread on glass, and perhaps most importantly, are recruited to thrombi *in vivo* when injected into a mouse injury model[238].

Forward programming iPSC can be used to generate large numbers of mature megakaryocytes through the virally mediated transduction of key MK transcription factors. The benefits of forward programming approaches is that they typically yield significantly larger numbers of progenitors by overcoming the asynchronous nature of cell differentiation as driven by directed cytokine treatment[239]. In 2014 Nakamura *et al.* reported the generation of an immortalised megakaryocyte progenitor cell line (imMKCL) through the introduction of a doxycycline inducible cassette regulating the expression of BCL-XL, C-MYC, and BMI1[240]. The resulting cells express markers of mature MKs and are sustainable in culture for 5 months. Upon the removal of doxycyline, and the subsequent loss of expression in these transgenes, imMKCLs undergo terminal maturation and release mature platelets which respond to a host of agonists[240]. Critically however, these functional responses are substantially lower than those evidenced by platelets isolated from human blood.

Similarly Moreau *et al.* over expressed FLI1, GATA1, and TAL1 to generate iPSC-MKs

in large numbers. These cells remain viable in culture while expressing key markers (e.g. CD41, CD42) for up to 90 days, and are capable of producing small numbers of platelets. The authors report the generation of $2x10^{11}$ from 10^6 input iPSC, releasing a total of $1x10^{12}$ platelets - averaging just over half a platelet per cell [70].

These approaches all share similar shortcomings in that none of the cells generated closely resemble human bone marrow megakaryocytes. Regardless of the means of generation, the MKs produced are typically substantially smaller than adult, mature MKs in human bone marrow, and produce nowhere near the vast number of platelets needed to maintain haemostasis in a human. Similarly, the platelets produced resemble human circulating platelets in size and morphology, but respond poorly to agonists. While the authors of these various studies demonstrate recruitment of their artificially generated platelets to thrombi *in vivo* using mouse models, typically this will be substantially less than the level of recruitment evidenced by platelets derived from peripheral blood.

The development of bioreactors as a means by which to drive improved platelet production in these cells is one of a number of routes by which iPSC-MKs can be improved in terms of their platelet production capacity, however as of yet no bioreactor has successfully yielded the number of platelets per input MK demonstrated by bone marrow cells[241, 233]. A significant concern when using forward programmed cells is the potential impact of over-expressing tumuorogenic genes - of which C-MYC and FLI1 are prime examples.

The applications of iPSC derived MKs in disease modelling and functional, phenotypic screening hinges on their ability to produce proplatelets *in vitro*. Unfortunately to date only a handful of studies apply these methods as a means by which to interrogate the effects of, for example, particular patient mutations. This particular aspect of the field is discussed in section 3.1.3

3.1.2 Genome editing through CRISPR-Cas9

The discovery of an adaptive prokaryotic immune system referred to as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)- Cas (CRISPR associated) has led to what is arguably one of the most significant advances in modern biology. In 2012 this bacterial immune system was shown by Jinek *et al.* to be programmable for site specific DNA cleavage *in vitro*, a finding which was shortly followed by the application of this technique for genome editing in human cell culture[242, 243, 244]. While previous nuclease based gene editing approaches, most notably Zinc Fingers and TALENs (Transcription Activator-Like Effector Nuclease), brought the possibility of site-specific genome editing to the study of gene function, issues with complexity and cost have always hampered efforts to make either technique widespread and readily available[245]. CRISPR-Cas allows for inexpensive, efficient, and relatively simple site specific genome editing both *in vitro* and *in vivo*.

The key components of CRISPR-Cas systems in mammalian genome editing are an endonuclease (most often the Cas9 protein from *S. pyogenes*) and a single guide RNA (sgRNA) which targets the protein to a sequence of interest (Figure 3.1)[242, 243, 244]. Cas endonucleases must bind an sgRNA targeting a region of genomic DNA immediately adjacent to a Protospacer Adjacent Motif (PAM) site (Figure 3.1)[246]. This is specific to the species from which the Cas endonuclease is derived. In the most commonly used protein, Cas9 from *S. pyogenes*, the PAM sequence is *NGG*[246]. Due to the frequency at which this particular PAM sequence is found in the mammalian genome, the *S. pyogenes* Cas9 based system has become extremely popular.

Upon the introduction of a Cas endonuclease and an sgRNA to which it can bind, whether through plasmid expression or transfection of an RNP (Ribonucleoprotein) complex, site specific double stranded breaks are induced immediately 3-4 base pairs 5' of the PAM[246, 247]. A number of Cas9 mutants have been developed, each of which feature mutations which alter the binding or cutting ability of the endonuclease. These include 'nickase' enzymes which

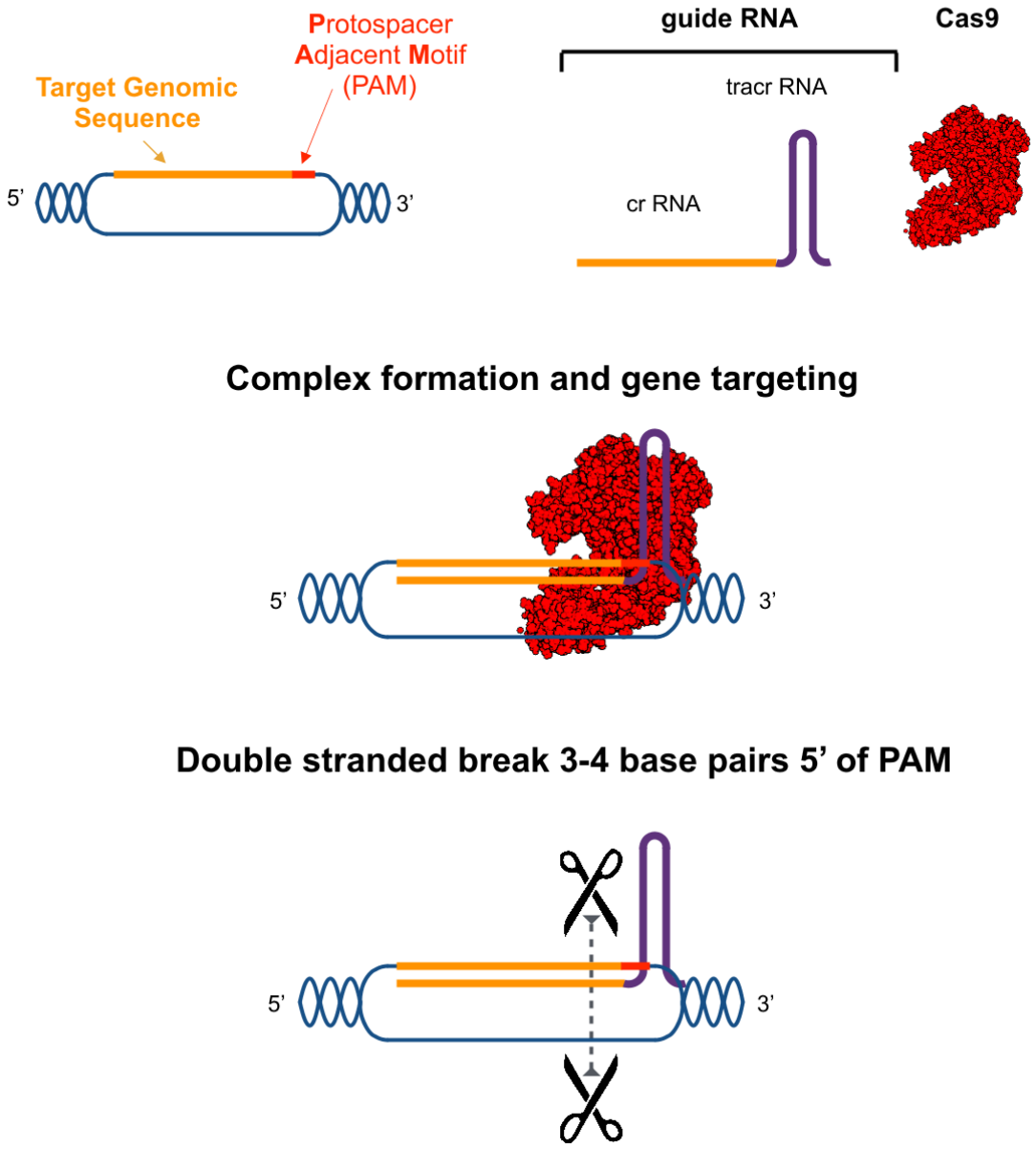

Figure 3.1: **Components of a CRISPR-Cas9 editing system.** Programmable CRISPR systems designed for mammalian gene editing rely on relatively basic components. A Cas endonuclease, most often the *S. pyogenes* Cas9 is expressed alongside a single guide RNA (sgRNA) targeting a specific genomic sequence immediately adjacent to a PAM site. sgRNAs are comprised of a CRISPR-RNA (crRNA) which is specific to the genomic target, and a tracrRNA which acts as a scaffold through which the Cas9 endonuclease binds the crRNA. On binding to an sgRNA, the Cas9 protein is able to bind to its adjacent PAM site and induce a double stranded break 3-4 base pairs immediately 5' of the PAM.

inactivate one of the two active sites of the enzymes to generate single stranded breaks (the D10A mutant for example), as well as inactivated or 'dead' Cas9s (dCas9) which retain their targeting ability but are unable to cleave DNA[246, 247, 248].

Nuclease directed genome editing ultimately depends on exploiting endogenous repair mechanisms following DNA damage, in this instance, a double stranded break. Once Cas9 has performed its cutting function, it falls to the cell to pursue either non-homologous end joining (NHEJ) or homology directed repair (HDR), the two key pathways for DNA damage repair (Figure 3.2). NHEJ is an error-prone pathway which will often result in indel mutations causing a premature stop codon which, if placed sufficiently towards the N-terminal coding region of the protein, can knock-out protein expression in the treated cell (Figure 3.2). HDR is a higher fidelity pathway which relies on the presence of a repair template (Figure 3.2). Its incidence is comparatively lower than NHEJ, however this particular mechanism can be exploited by co-transfecting with single stranded oligonucleotide or plasmid donors with sufficient homology to the target region for HDR. This hijacking of the DNA repair system is where the true potential of CRISPR-Cas lies, allowing for highly specific, targeted changes to the genome both *in vitro* and *in vivo*.

Microhomology mediated end joining (MMEJ) is another repair mechanism with potentially important implications for CRISPR-Cas9 experiments (Figure 3.2)[249]. This repair pathway occurs at frequencies higher than HDR, but lower than NHEJ, and can allow for targeted deletions as well as the insertion of larger sequences (e.g. fluorophores or resistance cassettes)[249].

The application of CRISPR-Cas 9 to achieve either targeted knock-outs or insertions in a relatively high throughput manner has had an undeniably significant impact on modern biology. In just 5 years the technology has developed in leaps and bounds, with a huge range of applications both *in vitro* and *in vivo*. In the following section, the relevance of these approaches to the study of inherited thrombocytopenias are discussed.

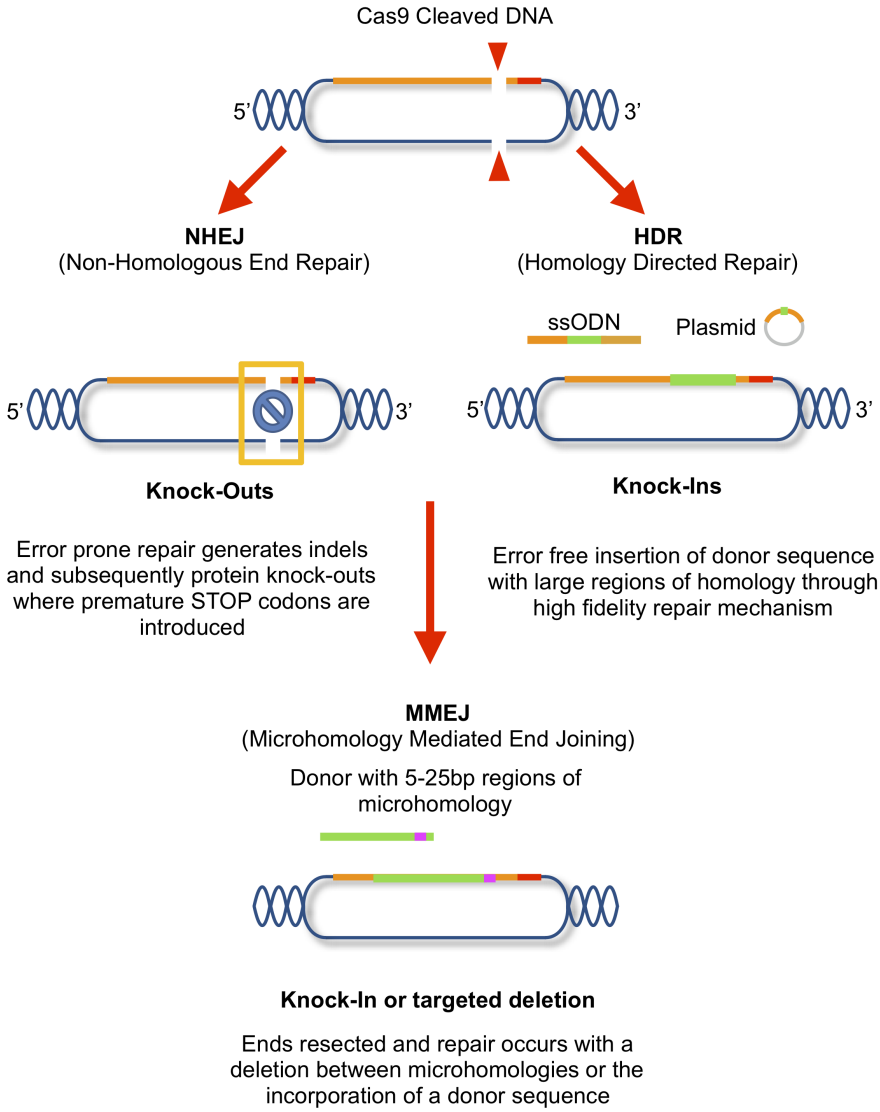

Figure 3.2: **Resolution of Cas9 induced double stranded breaks through endogenous repair mechanisms.** The CRISPR Cas system causes double stranded breaks which are resolved through endogenous repair mechanisms to introduce changes. In the absence of a designed repair template, NHEJ will constitute a large percentage of double strand break repairs. This is an error prone mechanism by which indels, and as a result, frameshifts and stop codons can be obtained. To introduce single base pair changes or large insertions like fluorophores, the high fidelity HDR route is often used. In this instance, a repair template (often a single stranded oligonucleotide for SNPs and a plasmid for large insertions) with significant regions of homology are introduced alongside the CRISPR-Cas9 components. HDR allows for the incorporation of these donor sequence in a relatively error-free manner. Finally microhomology mediated end joining is another means by which DSBs are resolved. This repair mechanism requires donors with regions of microhomology (5-25 bps) with the target strand. The donor DNA is resected until the region of microhomology is reached, at which point this is incorporated into the target DNA, most often resulting in the formation of large deletions.

3.1.3 Applying CRISPR-Cas9 genome editing and iPSC technologies to the study of inherited thrombocytopenia

Both technologies introduced in this chapter offer the potential to address long standing issues in the field of MK and platelet biology. In the absence of proplatelet producing cell lines, or an ability to culture large enough numbers of primary human MKs, iPSC derived megakaryocytes offer a potentially robust platform for *in vitro* experiments. Similarly, the development of CRISPR-Cas9 as a relatively simple method of gene editing allows for a means by which to edit genes of interest without the issues associated with over-expression (and practical issues associated with transfection in iPSC).

To date only a handful of studies apply these methods to the study of platelet production defects. This is at least partly due to the fact that iPSC-MK generation methods thus far have only demonstrated a limited capacity for proplatelet formation *in vitro*. A key criticism of many of the methods discussed in this chapter is the lack of quality imaging or immunofluorescence in iPSC derived MKs, and as such there is no indication from the literature that the use of iPSC-MKs is a viable means by which to assay platelet production *in vitro*.

The bulk of current iPSC-MK based studies to date use patient derived tissue to generate stem cell lines for phenotypic study. Notable examples include studies by multiple groups which reprogram cells from patients with a familial platelet disorder. In these works a loss of MK differentiation is reported recapitulating the thrombocytopenia reported in the patients involved[250, 251, 252]. Similar work has been performed in stem cells derived from patients with Paris-Trousseau syndrome and Glanzmann thrombasthenia[253, 239]. Critically proplatelet formation is not assayed in any of these works, nor are images of the iPSC-MKs reported.

Kahr *et al.* recently published an investigation into the loss of a haematopoeitic specific component of the Arp2/3 complex - *ARPC1B*. In this elegant study, two unrelated patients with a loss of function in *ArRPC1B* present with similar haematopoeitic deficiencies, including

eosinophilia, inflammatory disease and microthrombocytopenia[163]. The authors establish the phenotype observed using samples obtained from both patients, but also go on to use CRISPR-Cas9 to delete the *ARPC1B* complex in the imMKCL line from Namakura *et al.*. They proceed to report a reduction in proplatelet formation in edited cells versus wild type using quantitative SIM.

A substantial criticism of this latter part of the work is the appearance of the 'proplatelet forming' imMKCLs reported. Both cells are approximately 10μm in size, with no polyploidy evident in either of the cells shown. The proplatelet extensions reported do not resemble the beaded structures evidenced by foetal liver MKs, with no clearly defined tubulin rings which are considered hallmarks of proplatelet structures[163].

Reprogramming patient material is one of two methods of stem cell based disease modelling. Stem cell lines derived from patient tissue can successfully recapitulate the disease phenotype observed, and provide a powerful means by which to interrogate mechanistic questions - however, such an approach is less useful when dissecting the cause of a particular disease phenotype. Where multiple confounding mutations may be causative, or where patient sample is unavailable, being able to introduce particular mutations of interest into a genomic 'blank canvas' is a powerful method by which to specifically interrogate the effects of a novel mutation. CRISPR-Cas9 makes the latter method of disease modelling a viable option, as is discussed in the following section.

3.2 Aims

This chapter introduction has described significant advances in both genome editing ,and the application of stem cell technologies in the study of MK and platelet defects. To date, the utility of both these techniques to functionally characterise mutations has only been demonstrated by Kahr *et al.* in their study of the *ARPC1B* subunit. As detailed above, this study suffers from the same issues currently plaguing the application of iPSC-MK

differentiation for phenotypic screening - namely the lack of a method by which to generate large numbers of proplatelet forming cells *in vitro*.

This work aims to develop protocols by which to develop mature, proplatelet forming MKs from iPSC cultures, and to apply CRISPR-Cas9 genome editing to stem cells to replicate patient mutations and tag proteins of interest for microscopic study and analysis. The aim is to achieve a work flow demonstrated in figure 3.3, whereby candidate mutations can be efficiently modelled in an iPSC-MK model to evaluate the effects of said mutation on platelet production. This would be a substantial advance in the field, and offers a powerful *in vitro* tool in a field which has suffered from the lack of such an approach.

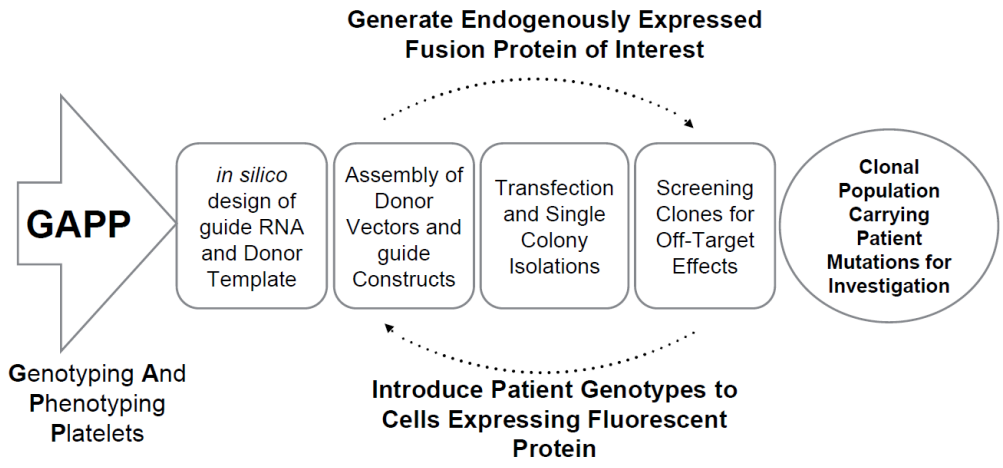

Figure 3.3: **Design of work flow aiming to apply CRISPR to the study of inherited thrombocytopenia.**Example of workflow based on output from the Genotyping and Phenotyping of Platelets (GAPP) study. Unique patient mutations are first identified by the GAPP project, and proteins of interest identified in this manner are placed in a CRISPR-Cas pipeline to first generate fluorescently tagged fusion proteins expressed at near endogenous levels. Patient specific mutations are then introduced to generate clonal populations expressing both wild type fusion proteins as controls, and mutant proteins for investigation into novel molecular causes of thrombocytopenia.

3.3 Materials and Methods

3.3.1 Stem Cell Culture

Gibco human episomal induced pluripotent stem cell line was purchased from Thermo Scientific and cultured on Geltrex basement membrane in StemFlex medium (Thermo Scientific).

Routine passaging was performed using EDTA (Sigma), with single cell seeding performed for transfection and attempted clonal isolation through the use of TryplE (Thermo Scientific). Briefly, cells were washed twice with PBS and once with either EDTA (for clump passaging) or TrypLE (for single cell) before incubation in 1mL relevant detachment media for 3 minutes at 37°C.

For clump passaging, EDTA was removed and 1mL of StemFlex added. Cells were detached by triurtating media onto the bottom of the well and subsequently adding the required volume to fresh media (in a new, GelTrex coated plate).

For single cell seeding, TrypLE was diluted in 2mL StemFlex and the solution added to a 15mL falcon tube for centrifugation at 200g for 4 minutes. The supernatant was then discarded and the cell pellet resuspended in the required volume.

3.3.2 iPSC MK differentiation

iPSC differentiation to mature, proplatelet forming megakaryocytes is discussed in the results section of this chapter, and is modified based on an original protocol published by Feng *et al.*[238].

To summarise, cells were detached by the clump passaging protocol described in section 3.3.1 and seeded on dishes coated with Collagen Type IV (Advanced Biomatrix) at $5\mu g/cm^2$. Cells were seeded overnight with RevitaCell to support survival on the new basement substrate.

To begin the protocol cells were washed twice and incubated in phase I medium comprised of APELII medium (Stem Cell Technologies) supplemented with BMP-4 (Thermo Scientific),

FGF-β, and VEGF (Stem Cell Technologies) at 50ng/mL each. Cells were incubated at 5% oxygen for the first four days of the protocol before being placed in a standard cell culture incubator for a further two days in freshly made phase I medium.

At day 6 of the protocol cells were incubated in phase II media comprised of StemSpan II, TPO (25 ng/mL), SCF (25ng/mL), Flt-3 (25ng/mL), Interleukin-3 (10ng/mL), Interleukin-6 (10ng/mL) and Heparin (5 U/mL) (all supplied by Stem Cell Technologies).

Each day in phase II media suspension cells were spun down at 400g for 5 minutes and frozen in 10% FBS/DMSO. After 5 days of collection, all frozen cells were thawed for terminal differentiation.

Terminal differentiation was performed by incubating cells in StemSpan II with heparin (5U/mL) and Stem Cell Technologies Megakaryocyte Expansion supplement on low attachment dishes (Corning).

3.3.3 FACS antibodies

Validation of surface marker expression was performed on an Accuri B6 as described in previous chapters. Cells were labelled by either first detaching (if cells were adherent) with TrypLE (Thermo Scientific) before centrifugation at 400 x g for 5 minutes. The pellet was then re-suspended in FACS buffer, comprised of PBS, 2mM EDTA, and 2% FBS. The relevant conjugated antibodies were added at a dilution of 2:50 and samples were incubated on ice for 20 minutes before the addition of 1mL PBS, and 2 subsequent washing and spinning steps. Finally, cells were re-suspended in 100μL FACS buffer before analysis. Undifferentiated iPSC were similarly treated to generate control samples.

The following antibodies were used to determine iPSC differentiation in this chapter: CD31 Alexa 488 conjugate (Thermo Scientific), CD42b APC conjugate (BD Pharmingen), CD34 Alexa Fluor 488 conjugate (Thermo Scientific), CD41 Alexa Fluor 488 conjugate (Thermo Scientific), and CD144 APC conjugate (BD Pharmingen).

3.3.4 RNP Complexes

The IDT Alt-R®RNP system was used for the relevant experiments in this chapter. crRNAs were ordered at 2nmol (following the design of *TubA1B* targeting guide C reported in Chapter 2) and resuspended in 20μL TE buffer (IDT) for a final concentration of 100μM. Atto-555 labelled tracrRNAs were ordered at 5nmol and resuspended in a volume of 50μL for a final concentration of 100μM.

To prepare small guide RNAs (sgRNA), equimolar ratios of both crRNA and tracrRNA were mixed with Nuclease Free Duplex Buffer (IDT). This mix was then incubated at 95°C for 5 minutes before allowing the reaction mix to cool at -1°C/second to 25°C. This mix was then spun down and complexed with either HiFi Cas9 V3 (1081058 - IDT) or Cas9 R10A nickase (1081063 - IDT) purified Alt-R®.

Cas9 protein was diluted to 6μg per transfection and incubated with an equal volume of annealed sgRNA. This mix was left for 30 minutes at room temperature to form complete and stable RNP complexes.

3.3.5 Stem Cell Transfection

iPSC transfection was performed using Lipofectamine Stem according to manufacturer instructions. Briefly, iPSC were seeded on 24 well dishes coated with Geltrex at 50,000 cells per well. After an overnight incubation in StemFlex with RevitaCell, cells were washed twice with PBS and once with OptiMem before incubation in OptiMem with RevitaCell.

RNP complexes were prepared as described in section 3.3.4 and resuspended in 25μL OptiMem per reaction. A Lipofectamine Stem master mix was prepared using 25μL OptiMem and 2μL Lipofectamine STEM per reaction (4μL if a donor template is included). Equal volumes of both Lipofectamine and RNP mix were incubated to form lipofection complexes over a 10 minute incubation at room temperature. The final transfection mix was added to cells in OptiMem and left for 4 hours before the addition of StemFlex medium (and any

relevant small molecules).

Measurement of iPSC transfection efficiency after treatment with Lipofectamine STEM and IDT RNP complexes was performed using manual cell counting in Evos acquired images (Phase contrast and fluorescence).

3.3.6 Small molecule targeting of the HDR pathway

Small molecules were purchased and resuspended in DMSO before dilution to the required concentrations. SCR-7, RS-1, (Sigma), NU7026 (Sigma), Trichostatin A (Sigma), MLN4924 (Adooq BioScience), NSC 15520 (ChemBridge).

3.4 Results

3.4.1 Establishment of an iPSC Culture System

A plethora of iPSC culture systems are currently available, usually comprised of different basement membranes or feeder layers, cytokines, or nutritional supplements. The choice of system ultimately has a significant impact on downstream applications. For our purposes, a culture system which supports cells when 'challenged' through transfection or single cell cloning, without detrimental effects on potency, survival, or karyotype, would be key for successful genome editing and differentiation.

To that end a range of different feeder free basement membranes and media were trialled, including matrigel and vitronectin for the former, Essential 8 and NutriStem for the latter (data not shown). Ultimately however, we found that growing our iPSCs on GelTrex basement membrane and StemFlex media supported growth with virtually no differentiation and minimal cell death. As shown in Figure 3.4A, even after defrosting, cultures show minimal cell death with tight, clearly defined colonies lacking the spread differentiated cells typical of iPSC post thaw.

Despite freezing and subsequent thaw and culture, cells maintain a normal, healthy karyotype (Figure 3.4(B)). In our hands this culture system is ideal for our subsequent treatment of the cells.

3.4.2 Electroporation of iPSC with Cas9 and guide expressing plasmid

iPSC transfection was, at the start of this project, extremely difficult in the absence of viral transduction. Initial work trialled a number of lipofection agents (Lipofectamine 3000, Fugene HD, TransIT) and electroporation methods (Neon electroporation, Amaxa Nucleofection) to introduce the px459 guide C expressing plasmid and mEGFP *TubA1B* donor described in

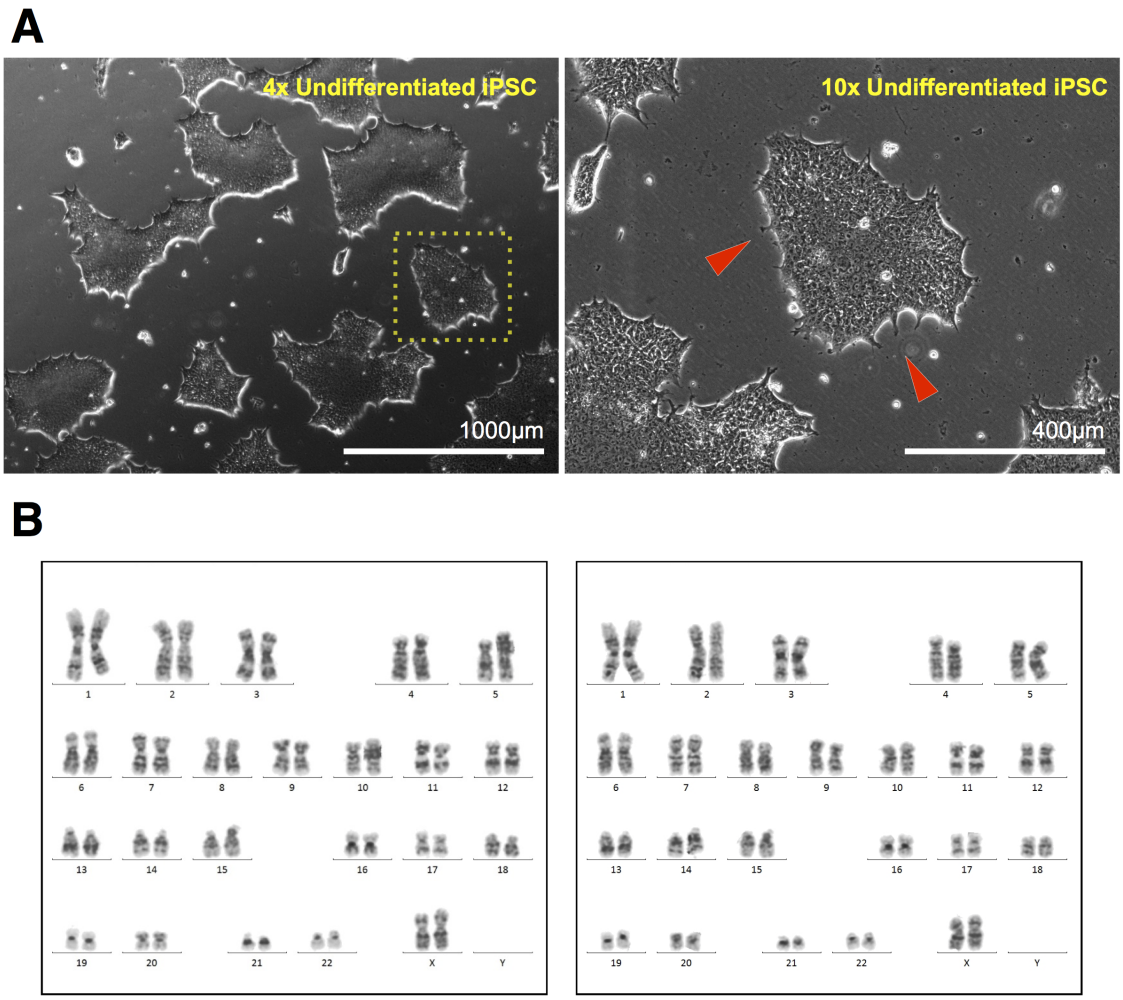

Figure 3.4: **Establishment of a routine iPSC culture system with a stable karyotype.** (A) After trialling a range of basement membranes and culture media, Gibco iPSC StemFlex media and GelTrex were found to allow for routine culture of iPSC without unwanted differentiation around colony edges. Images of a culture after defrosting, an event which usually drives differentiation and cell death, show a stable culture with no notable differentiation around colony edges at both 4x and 10x magnification. Clean colony edges are observed (indicated by red arrows). (B) Cells maintain a normal, healthy female karyotype as indicated by two representative images. (Fixed samples were sent to Birmingham Womens Hospital for karyotyping).

chapter 2 of this thesis. The aim of this was to trial an established method of using CRISPR to evaluate the efficiency of HDR and fluorophore integration in iPSC.

Of these early attempts at replicating the cell line data, Neon electroporation was the only successful method (Figure 3.5), generating mEGFP positive cells at 0.1% efficiency (compared to as much as 10% in cell lines) in cells electroporated at 1200V and with a 30ms pulse width. Other settings were not successful in introducing the mEGFP tag.

Unfortunately this efficiency, while consistent with the HDR rates observed in the literature, is extremely low. When targeting MK and platelet specific genes which are not expressed in undifferentiated iPSC it would be almost impossible to isolate the 0.1% of the population carrying an insertion.

3.4.3 Transfection with Lipofectamine STEM and RNP complexes

Thus far while small molecule treatment does have a positive effect on tag integration in both cell lines and iPSC, the effects observed have not achieved the level of integration needed to make routine gene tagging in stem cells a viable option. While this is consistent with much of the literature, it remains a substantial obstacle to this work.

With this in mind both the method of transfection and Cas9 delivery were altered as per recently published work which shows an increase in cleavage efficiency and HDR through the delivery of ribonucleoprotein complexes (RNPs)[254]. Jacobi *et al.* demonstrate an improvement in CRISPR efficiencies when purified Cas9 protein is complexed with a synthetic sgRNA and subsequently electroporated or lipofected into cells.

Firstly the method of transfection was altered as a sample of the newly developed Lipofectamine STEM was kindly provided by Thermo Scientific. As demonstrated in figure 3.6, the use of Lipofectamine STEM with IDT's RNP complex allows for efficient transfection into iPSC. The IDT Alt-R®system is comprised of a user designed guide element which is annealed to a fluorescently labelled tracrRNA before complexing with purified Cas9 protein.

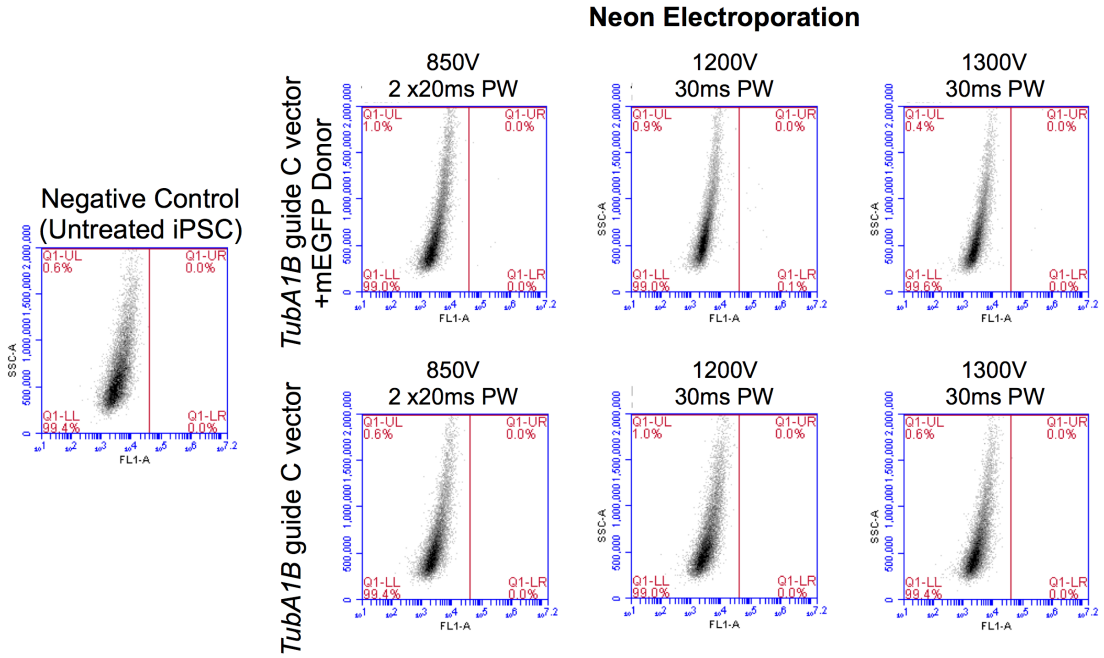

Figure 3.5: **Electroporation of Cas9 and guide expressing plasmid targeting TubA1B does not achieve efficient integration of donor sequence.** iPSC were electroporated using the Neon electroporation system and an optimised protocoll provided by the manufacturer for Gibco iPSC. Cells were mixed with the px459 plasmid containing theTubA1B targeting guide C used in Chapter with and without the relevant mEGFP donor. 3 different electroporation conditions were used as recommended by the manufacturer, with only one achieving an extremely low efficiency of integration (0.1% at 1200V and a pulse width of 30ms).

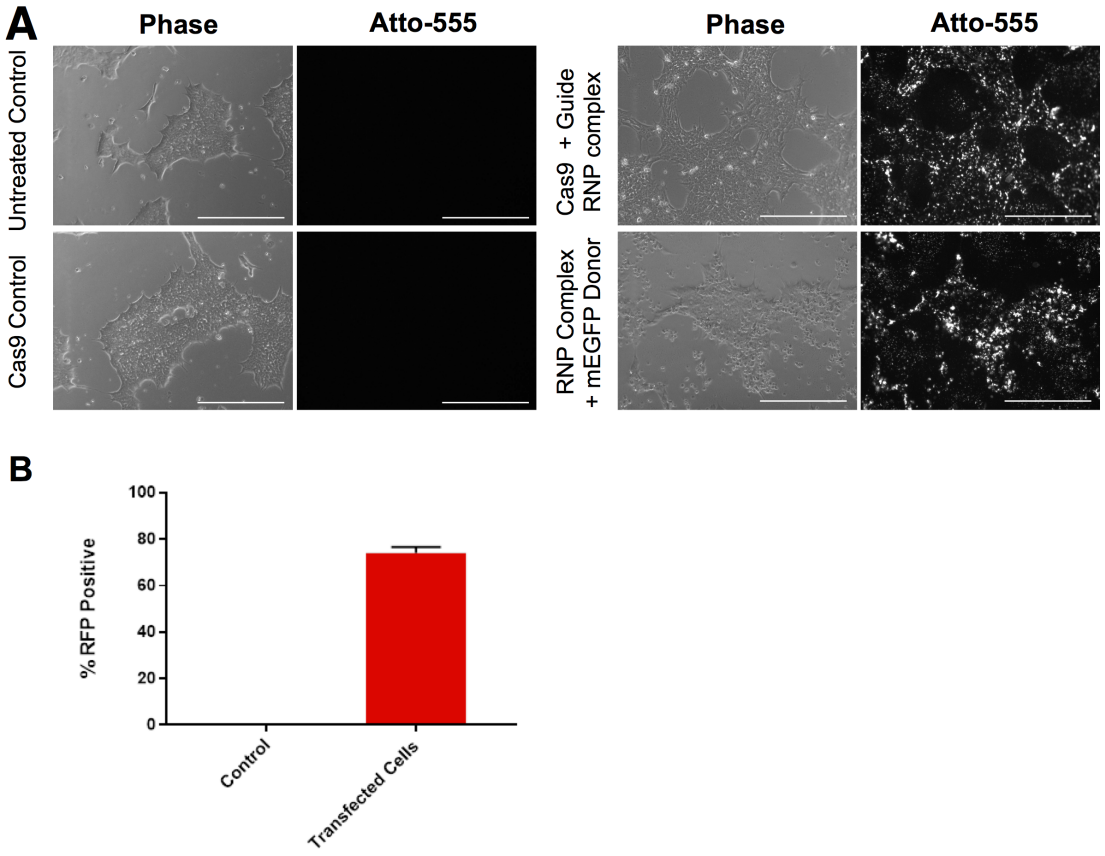

Figure 3.6: **Efficient iPSC transfection using Lipofectamine Stem.** A test batch of Lipofectamine STEM was provided for test transfections. We performed lipofections using this kit and IDT RNP complexes which include a tracrRNA labelled with Atto555. (A) iPSC were plated over night in ROCK inhibitor before lipofection in OptiMem. Cells were transfected with IDT RNP complexes and incubated in OptiMem with ROCK inhibitor for 4 hours before the addition of StemFlex culture medium. Untreated and Cas9 control only cells show no fluorescence using illumination through a 555 cube on an Evos-FL (10x objective, 400m scale bar), in contrast cells treated with an RNP complex and tracr show a high number of cells positive for signal, as do cells treated with both the guide complex and an mEGFP donor. (B) Analysis of transfected images show a transfection efficiency of 71.08 % ($\pm$1.28). n =3, S.D.

The fluorescent RNP complex is transfected and is easily visualised using a bench top Evos system for a convenient assessment of transfection efficiency.

Transfection efficiencies as high as 71.08% (±1.28)have not been reported through lipofection in iPSC, but perhaps more critically, cell survival and health as demonstrated in figure 3.6A remains excellent, while electroporation is notoriously harsh and causes significant cell death. In karyotypically sensitive cells like iPSC, this is vital to generating and maintaining healthy CRISPR edited lines.

With an efficient method of introducing a Cas9-guide complex, iPSC were then treated with the Alt-R®RNP complex carrying *TubA1B* targeting guide C and either double stranded or plasmid donor DNA. While plasmid donors worked extremely well for the cell line work reported in chapter 2, there is some debate as to whether it is the appropriate format for knock-in donor sequences in HDR focussed CRISPR experiments[247].

3.4.4 Testing small molecule enhancers of HDR

In 2016 a number of papers reported that small molecule inhibitors of NHEJ could increase the rate of fluorophore integration through HDR in CRISPR targeted cells. Of these, the most often published small molecule was the DNA Ligase IV inhibitor SCR7, with reports of increased HDR in cell lines, primary cells, and iPSC for both small and large insertions[255, 256, 257]. Of the many other small molecules reported, RS-1 showed a substantial improvement in HDR in cell lines and on insertion into mouse embryos - interestingly, the authors show a negligible effect of SCR7 in comparable experiments[258]. RS-1 has been previously reported as a stimulator of RAD51, a member of the RecA family mediating homology directed repair[259].

Based on these works, RS-1 and SCR7 were used at a range of concentrations recommended by the literature to determine the effects of these small molecules on the integration of *TubA1B* mEGFP in Hek293T. As shown in figure 3.7, treatment with both RS-1 and SCR-7 improves

the rate of HDR, with significant increases at an RS-1 dose of 7.5μM, and SCR-7 doses of 0.5μM and 1μM, both consistent with reports in the literature. Interestingly we do not observe a clear dose dependent increase in integration, though this is likely due to the deleterious effects of homozygous knock-in at the *TubA1B* locus.

As SCR7 at 1μM achieves the most substantial fold increase in HDR (1.763 fold $\pm$0.15), this concentration was taken forward for iPSC experiments. iPSC were transfected in an identical experiment and treated with either DMSO or SCR7. SCR7 treated cells demonstrated an increase in integration, however the rate of HDR remains extremely low (Figure 3.8). In SCR7 treated cells a total of 47 positive cells were found out of a total of 65,253. In the DMSO control 44 positive cells were found in a total of 172,922 (Figure 3.8B) (0.2%).

While the treatment reproduced a rate of HDR in iPSC consistent with the literature, the overall rate of fluorophore integration remains too low for our purposes. With SCR7 treatment we would still need to screen 1,388 individual stem cell cultures to find a single positive clone.

3.4.5 Testing Cas9 nickase and CRISPY inhibition

A recent publication described the rational inhibition of key components of NHEJ machinery, reporting up to 7.20 fold increases in integration in treated cells with approximately 20% of targeted chromosomes successfully edited[260]. These are the highest rates of HDR reported in the literature thus far, and so we acquired components of the CRISPY mix to determine whether CRISPY could improve our editing rates[260].

The CRISPY mix described by Riesenberg *et al.* refers to a cocktail of small molecules targeting specific components of the cellular repair machinery to inhibit NHEJ and improve HDR. NU7026 is a small molecule inhibitor of Ku 70/80, while other components of the mix are enhancers of particular components of the HDR machinery. MLN4924 and Trichostatin A are enhancers of CtiP and ATM, while NSC 15520 (which we were not able to procure) is

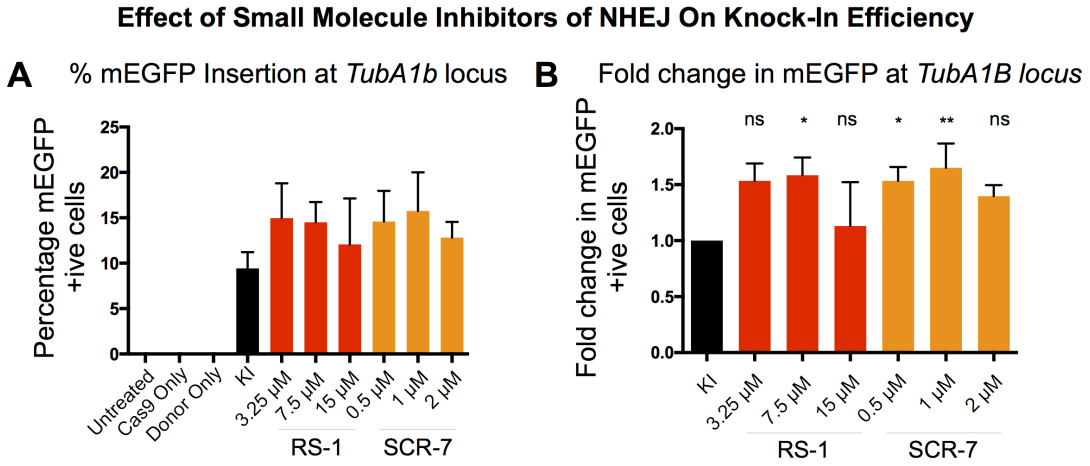

Figure 3.7: **Effect of Small Molecule Inhibitors of NHEJ On Knock-In Efficiency.** Hek293T cells were treated with Cas9 and the TubA1B targeting guide C described in (insert where here) in the presence and absence of an mEGFP donor and two NHEJ inhibitors. KI indicates knock-in controls, samples treated with the complete combination of vectors but no additional small molecules. (A) RS-1 and SCR-7. Treatment with these inhibitors increased the percentage of cells positive for an mEGFP insertion in all treated samples. (B) Normalisation to fold change in insertion reveals significant increases at 7.5μM of RS-1, and 0.5μM and 1μM of SCR-7, with the largest fold change observed in SCR-7 treated samples at 1μM with a 1.763 fold increase in insertion (SD 0.3198). (Data tested with a One Way ANOVA and multiple comparisons with no corrections n = 4).

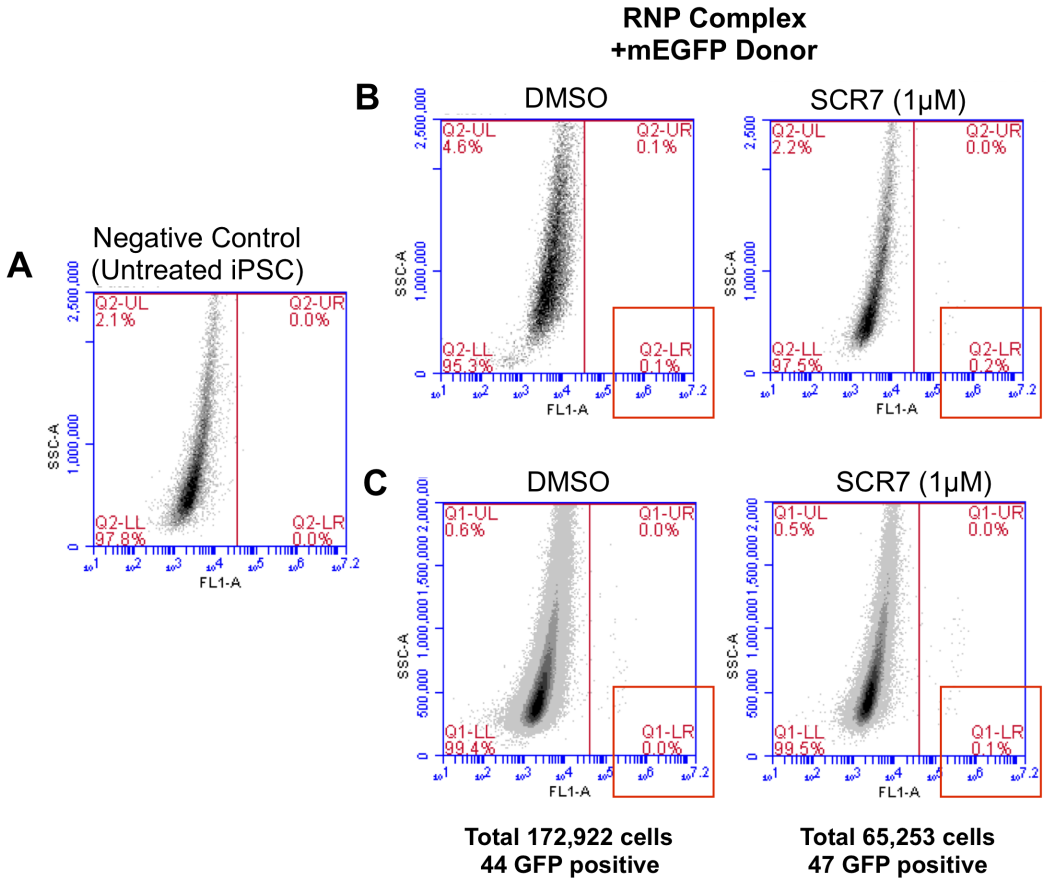

Figure 3.8: **Treatment with SCR7 increases rate of HDR.** To determine whether the increase in HDR observed through SCR7 treatment in Hek293T achieved a similar result in iPSC, stem cells were electroporated with an mEGFP donor, Cas9 and *TubA1B* targeting guide C in the presence of 2μM SCR7 or DMSO. (A) In untreated iPSC, no integration of the mEGFP tag is observed. (B) In DMSO treated cells 0.1% population is positive for mEGFP, while 0.2% of the SCR7 treated population is positive. This doubling of HDR is consistent with a similar treatment in Hek293T, however the overall rate of HDR is extremely low. (B) Samples were run to completion and a total of 44 positive cells are observed in the DMSO treated sample out of 172,922 - while 47 positive cells are observed in the SCR7 treated sample (out of 65,253).

thought to block p53 binding to RPA and RAD9. The authors show significant increases in iPSC editing through the introduction of both large cassettes (a BFP sequence) and single nucleotide changes[260]. This work demonstrates the highest reported improvement in HDR, with as many as 50% of chromosomes edited in particular targets.

The authors report a significant improvement when used in an iPSC line with an inducible Cas9 nickase cassette, and as such we adapted our RNP complex to include purified Cas9 D10A nickase protein. Cells were transfected with an mEGFP donor in either a dsDNA format (generated by PCR), or as a plasmid as in chapter 2. Unfortunately we found, as demonstrated in figure 3.9, that CRISPY treatment did not alter the rate of mEGFP insertion when targeting the *TubA1B* gene.

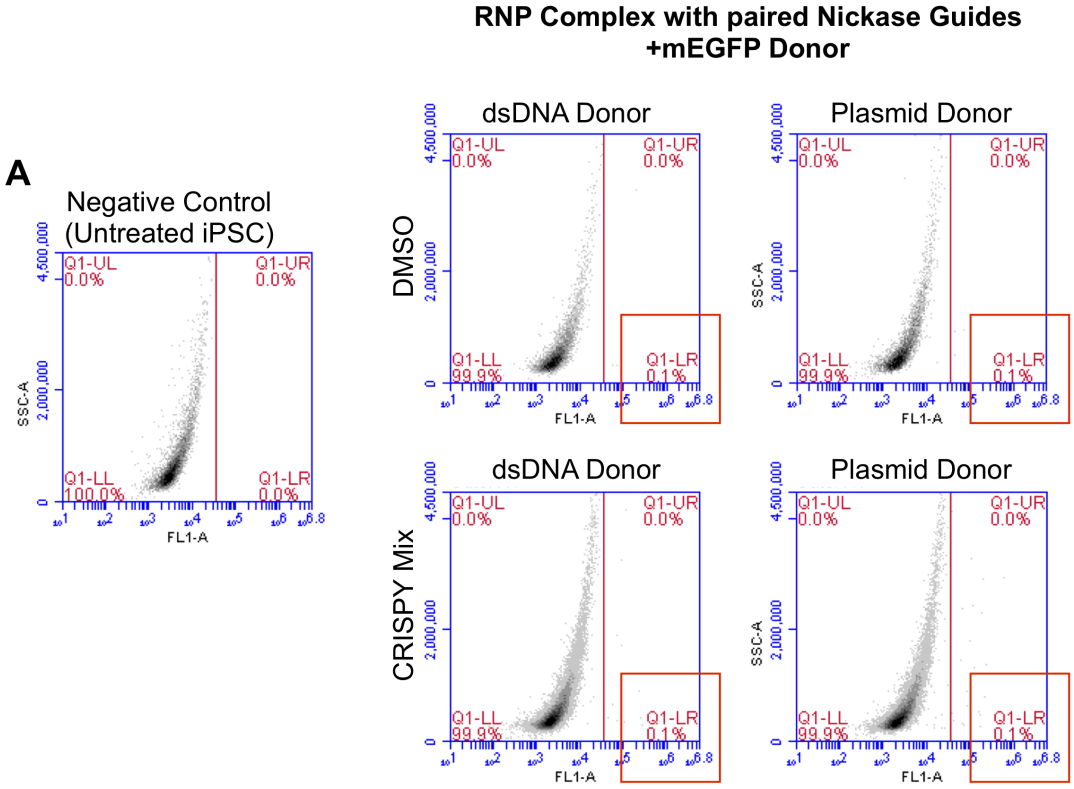

Figure 3.9: **Use of paired Cas9 nickase guides and CRISPY inhibition does not significantly increase HDR at *TubA1B*.** iPSC transfected with RNP complexes comprised of a Cas9 D10A mutant and two paired guide RNAs and an *TubA1B* mEGFP donor (either plasmid or dsDNA) were screened for HDR with FACS after 48 hours in the presence of either DMSO or CRISPY small molecule mix. (A) Untreated, negative iPSC were negative for any cells in the FL-1 channel. DMSO treated cells with either a dsDNA or plasmid donor both demonstrate extremely low rates of integration (0.1%), which are not significantly improved on treatment with CRISPY.

3.4.6 Directed differentiation of iPSC megakaryocytes

One of the key aims of this project is the development of a method of phenotypic screening through the differentiation of iPSC into mature, proplatelet producing MKs. While significant advances have been made in the generation of MKs from iPSC in recent years, the application of these methods for mechanistic studies has been limited. Developing a robust assay of proplatelet formation in human, stem cell derived MKs would be a substantial advance in the field, particularly in conjunction with modern gene editing techniques.

To that end, with a robust iPSC culture system in place, the Feng *et al.* 2014 method of directed differentiation was adapted to produce large cultures of mature cells. Briefly, this method involves the generation of haemogenic endothelial progenitors by seeding iPSC on collagen IV in the presence of cytokines that drive pluripotent cells towards a primitive mesoderm. To encourage, this cells are incubated in hypoxic (5% oxygen) conditions for the first phase of differentiation. After 6 days the culture medium is changed to one carrying cytokines known to favour haematopoeitic differentiation. For the following week cells are harvested each day and frozen in 10% DMSO before thawing and further culture in a terminal differentiation step.

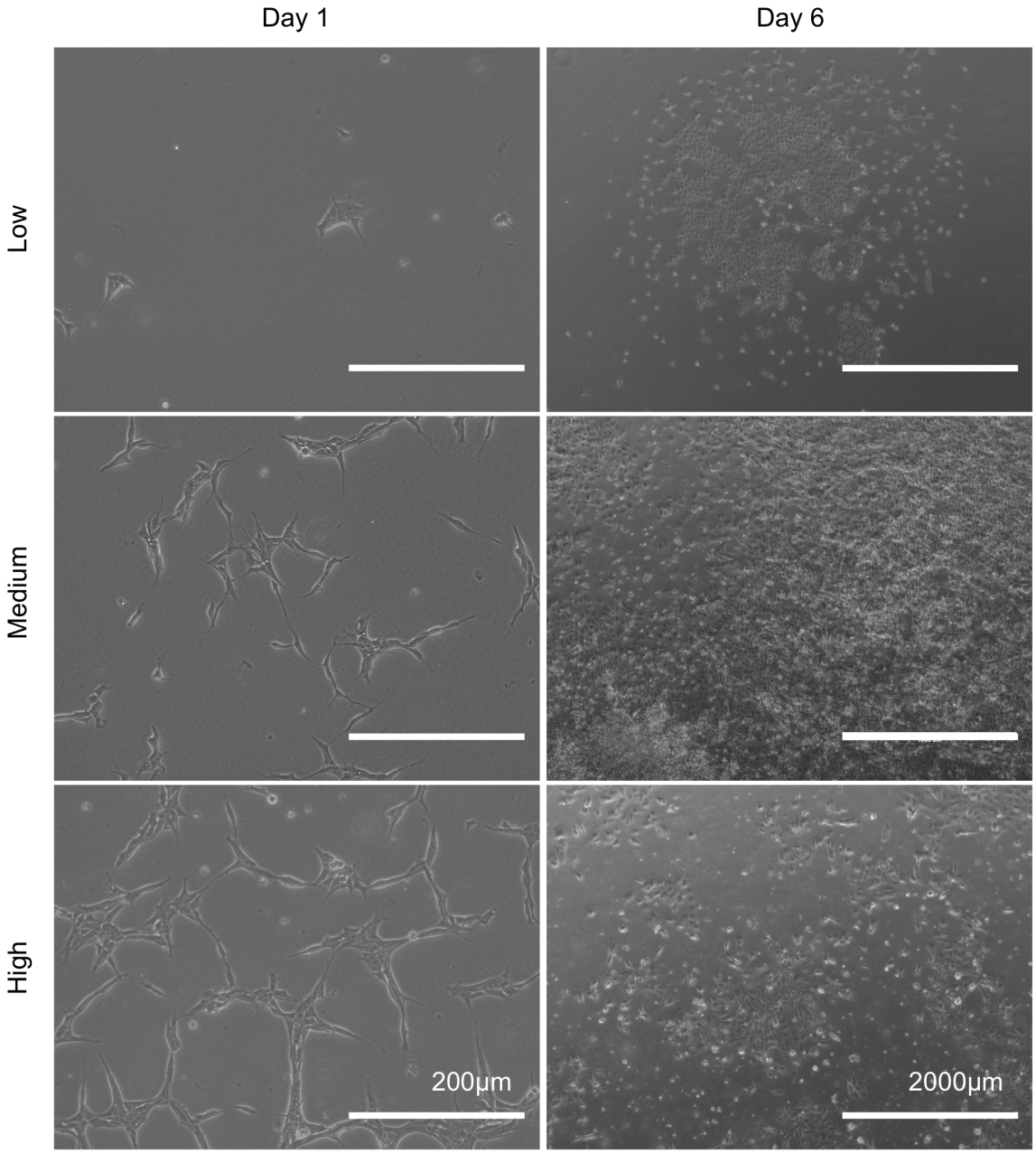

Figure 3.10: **Variation in confluence results in markedly different cultures.** Cells seeded on collagen IV at a low, medium, and high cell density demonstrate substantially different differentiated cultures on day 6 of this directed differentiation protocol. The low confluence sample demonstrates less outgrowth than its counter parts, while the higher confluence seems to result in increased cell death. The medium density results in a confluent monolayer with raised ridges in which cells appear to overlap with one another.

Early attempts to reproduce this protocol on a number of basement membranes and in different culture media proved ineffective. On matrigel and geltrex cells differentiate, but do not successfully generate haematopoeitc cells (data not shown). Similarly the formulation of the media originally used in the publication has since been altered, and we found that using these media similarly reduced the efficiency of haematopoeitic differentiation (data not shown). As such the original protocol has been altered as described in this results section.

Feng *et al.* do not report a seeding density in the original publication, and early experiments quickly revealed a significant impact of cell density on culture growth and architecture (Figure 3.10). We found cell survival on Collagen IV to be poor on single cell seeding, and as such proceeded with clump passaging to better preserve cell viability. Cells were seeded at a low (1/20 dilution), medium (1/10 dilution) and high (1/5) confluence of a routine passage (grown to approximately 75% confluence) and imaged as demonstrated in figure 3.10. The resulting population of haemogenic endothelium is dramatically different across cultures, with low seeding densities resulting in poorer out growth and more spread cells, while the higher cell density can achieve detrimental effects on cell health. At a 'medium' confluence, cells form a monolayer with distinctive ridges in which cells appear to overlap (Figure 3.10).

As the role of confluence had not been reported in the original paper, day 6 differentiated populations from each seeding density were screened by FACS to determine the expression of markers of haemogenic endothelium, and so which seeding density is best suited to the efficient generation of haematopoeitic progenitors. In this first phase of the differentiation protocol, we found APELII to be an optimal medium as suggested by a consultation with the manufacturer (Stem Cell Technologies). This base media was supplemented with 50ng/mL of BMP-4, FGF, and VEGF. Compared to unstained and undifferentiated iPSC controls, there is a substantial increase in key markers (CD31, CD34, and CD144) as shown in figure 3.11A. Cells seeded at a medium confluence demonstrate a substantial increase in both the percentage of CD34 and CD144 cells, with an increase in median fluorescence intensity in

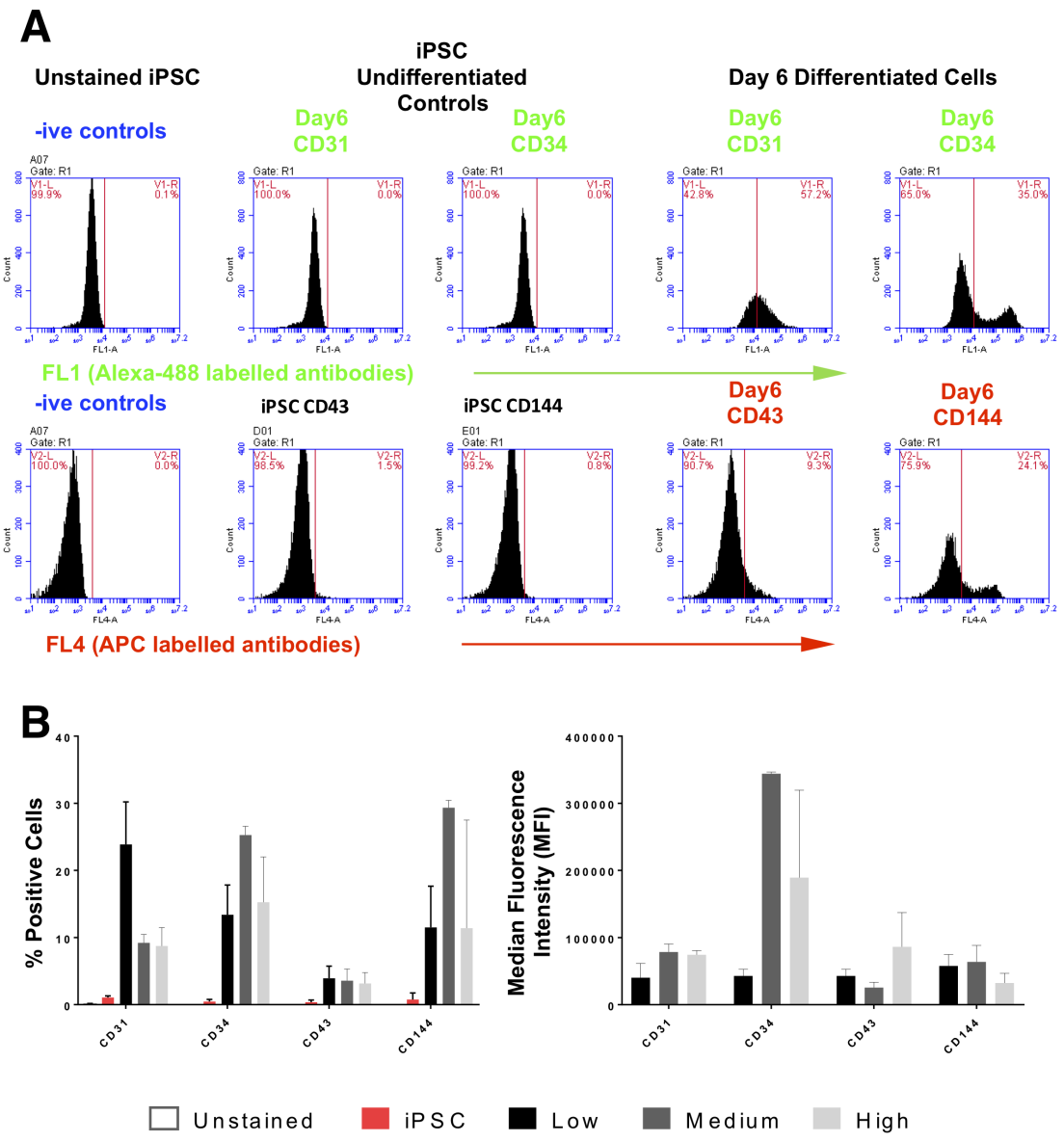

Figure 3.11: **Seeding density impacts the quality and number of haematopoeitic progenitors developed through directed differentiation.** iPSC differentiated following the first phase of the Feng protocol demonstrate a confluence dependent change in surface marker expression. Cells were stained with key markers of haemogenic endothelium, namely CD31, CD34 (both of which are detected in the FL-1 channel), CD144 and CD43 (which are detected in the FL-4 channel). (A) A notable increase in expression is observed in differentiated samples when compared to wild type and unstained cells. (B) More confluent cells appear to better express CD34, the key marker of haematopoeitic progenitors. While a lower confluence appears to favour the expression of CD31, a classic endothelial marker. Cells seeded at a higher density demonstrated more variance consistent with the higher cell death observed in these samples.

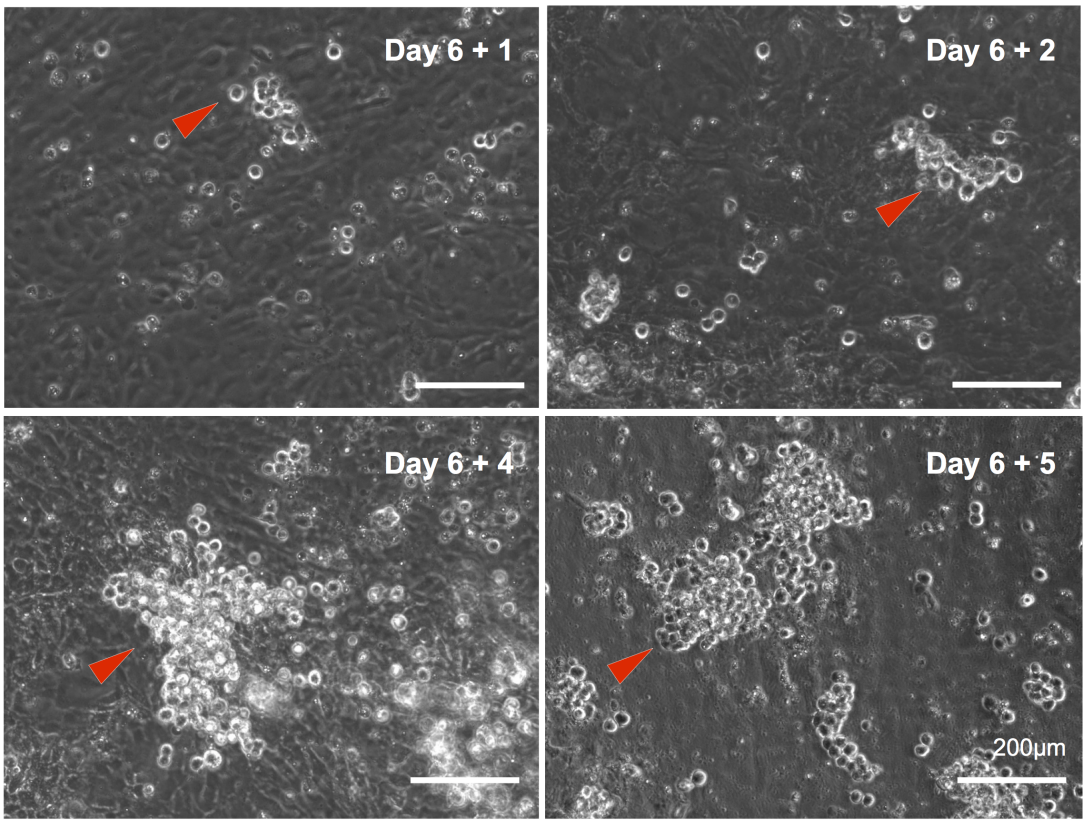

Figure 3.12: **Addition of haematopoeitic cytokines drives the development of MK progenitors.** The addition of phase II differentiation media results in the formation of progenitors and blast like cells (indicated by red arrows). APELII media supplemented with Flt-3, SCF, TPO (50ng/mL), IL-3, IL-6 (10ng/mL), and Heparin (5U/mL) drives the formation of these cells, which become evident at day 6+1 and day 6+2, albeit in substantially lower numbers. As the culture matures more progenitors form on day 6+4 and 6 +5.

CD34 also observed.

Interestingly, cells seeded at a lower density demonstrate an increase CD31 positive cells (figure 3.11B), suggesting that the lower density is better suited to the generation of endothelial like cells than haematopoeitic precursors. The variation evident in the high density seeded samples is likely due to the cell death observed as cells overgrow.

This data suggests a role for cell-cell signalling in generating a population of robust haematopoeitic progenitors. Interestingly as the images of the culture shows in figure 3.10, there are regions of overgrowth where cells appear to be self-organising into structures almost like embryoid bodies. Future work will interrogate these samples with immunfluorescence to determine which parts of this complex culture become haematopoeitic precursors.

With an optimal configuration for the first phase of this differentiation in hand, cultures were taken forward to the second phase in which haematopoeitic cytokines are introduced to induce the generation of haematopoeitic precursors and megakaryoblasts. As shown in figure 3.12, incubation in phase II media (comprised of APELII, 25ng each of Flt-3, SCF, TPO, 10ng each of IL-3, IL-6, and 5 units of heparin) results in the development of haematopoeitic precursors which resemble megakaryoblasts. The number of these cells increases each subsequent day, and to maintain the population at a rougly equal stage of maturation cells are harvested and frozen each day, with fresh media preparations added at days 3 and 6 of phase II.

Upon completion of the second phase of differentiation, frozen cells are thawed and resuspended in StemSpan II medium supplemented with StemCell Technology's Megakaryocyte Expansion Supplement and heparin at 5 U/mL. This MK supplement is comprised of TPO, SCF, IL-6, and IL-9, and has been shown by Feng *et al.* to facilitate the terminal differentiation of MK progenitors into mature megakaryocytes. Cells are grown on ultra-low attachment plates and imaged on day 1 and day 5 of terminal differentiation (Figure 3.13). At day 1 of the terminal differentiation, the cells observed are approximately 10μm in size

(Figure 3.13A). After 5 days of growth cells are notably larger at approximately 25μm in size, with some appearing substantially larger (Figure 3.13B).

Most methods of iPSC differentiation to the megakaryocytic lineage successfully generate blasts and mature MKs positive for CD41 and CD42. However the major impediment in the field has been moving these cells forward into a state of active platelet production both for the purposes of phenotypic screening and platelet generation for transfusion medicine.

This culture system yields cells which generate complex networks of proplatelet extensions very similar to those observed in murine foetal liver cells, the current gold-standard of *in vitro* platelet production (Figure 3.14). The appearance of this particular morphology is one of the key aims of this project, and to our knowledge is thus far the best example of iPSC-MK proplatelet formation reported.

Validation of this culture system is still underway, however early work shows that the consistent with markers and phenotypic features of mature iPSC-derived MKs (Figure 3.15). 60% of the terminally differentiated day 5 population are double positive for CD41 and CD42, compared to no such population in the undifferentiated iPSC control (Figure 3.15A). Immunofluorescence reveals polyploidy in analysed cells, with the characteristic lobed multiple nuclei of a mature megakaryocyte (Figure 3.15). Finally cells produce long, beaded extensions which, on tubulin staining, reveal beaded structures not dissimilar to those observed in proplatelet forming murine foetal liver cells.

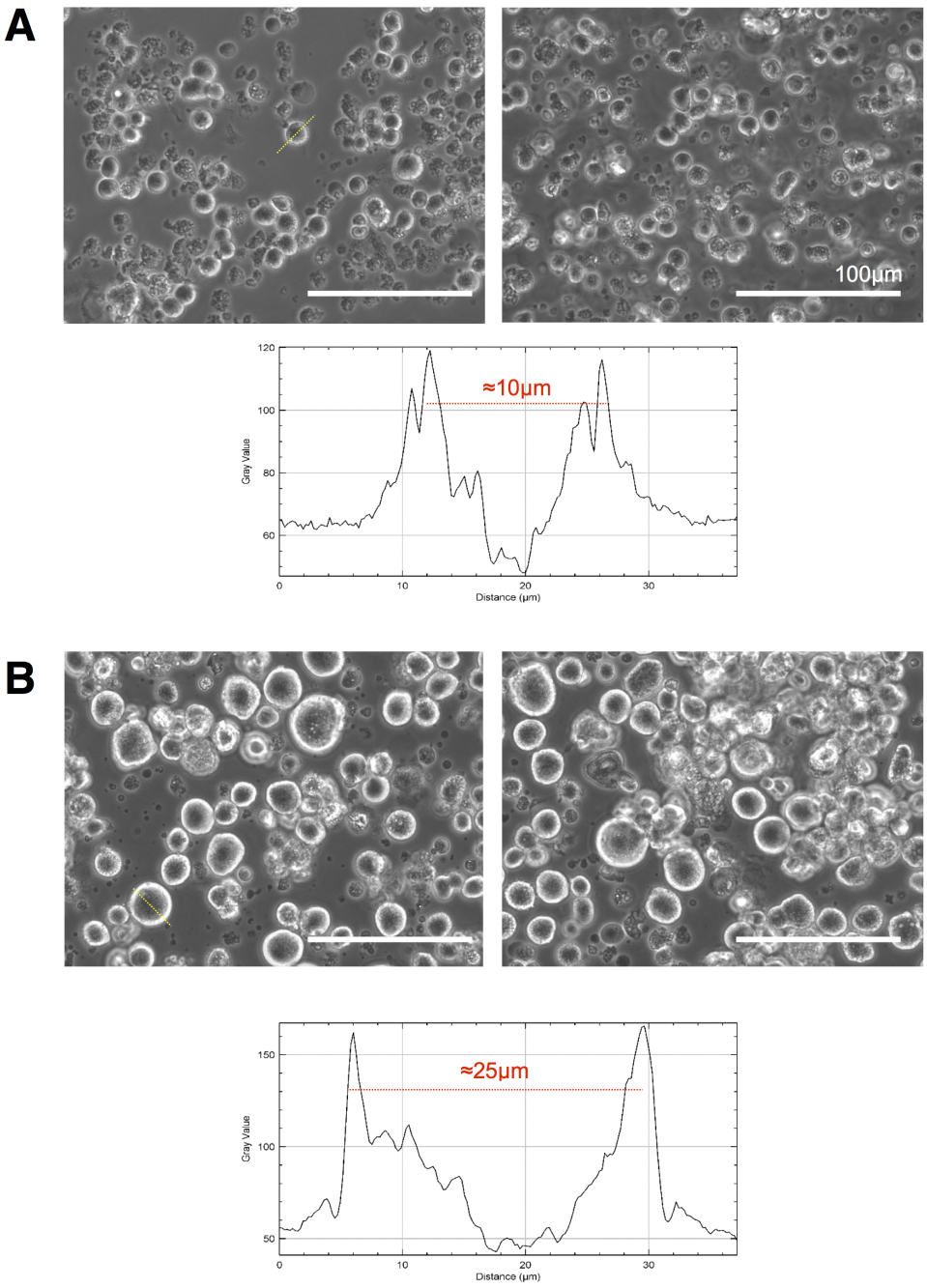

Figure 3.13: **Terminal differentiation of iPSC derived MKs.** On thawing and culturing in Megakaryocyte Expansion supplement, iPSC derived MKs demonstrate a marked increase in size preceeding proplatelet formation. (A) As indicated by representative images acquired on separate differentiations, on day one of the terminal differentiation protocol cells appear approximately 10μm in size. (B) By day 5 of the culture cells are substantially larger, typically 25μm in size. This increase in size is a hallmark of cultured MKs, and is consistent across experiments.

Proplatelet forming iPSC MK samples are also positive for platelet like particles (Figure 3.16A). These are small cells with a distinctive marginal band, no nucleus, and about 3μm in size. While the focus of the current protocol is the development of a proplatelet assay, the presence of these platelet like particles suggest that future work could study genetically altered platelets generated *ex vivo* through this experiment.

Figure 3.17 summarises the current version of the revised Feng protocol which has been developed in the last few months. The protocol differs in some ways from the original, primarily in the choice of media, density of cells, and preparation of samples in the final stages.

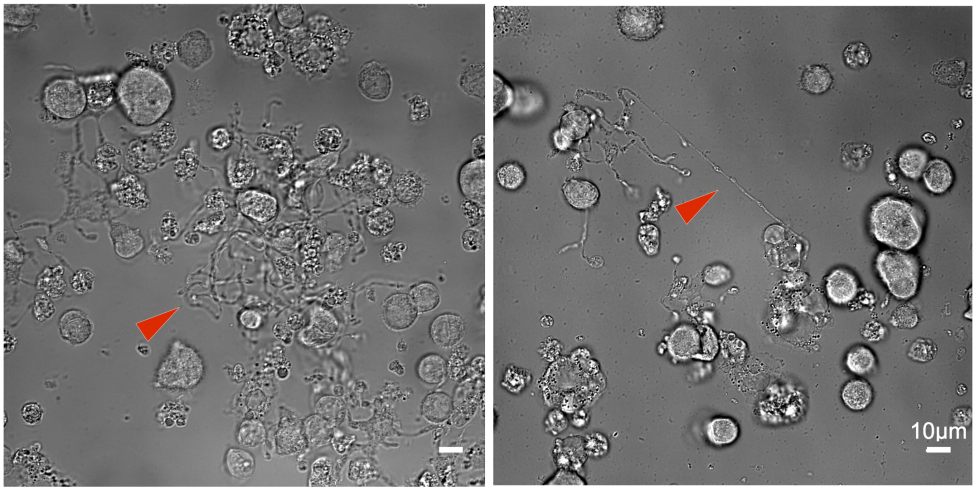

Figure 3.14: **Terminal differentiation of iPSC MKs results in the formation proplatelet forming cells.** Between day 4 and day 6 of the terminal differentiation of iPSC generated MKs cells demonstrate proplatelet production *in vitro*. Proplatelets appear as long, beaded extensions (indicated by red arrows) often covering up to 100μm fields of view, consistent with those observed in foetal liver murine cells.

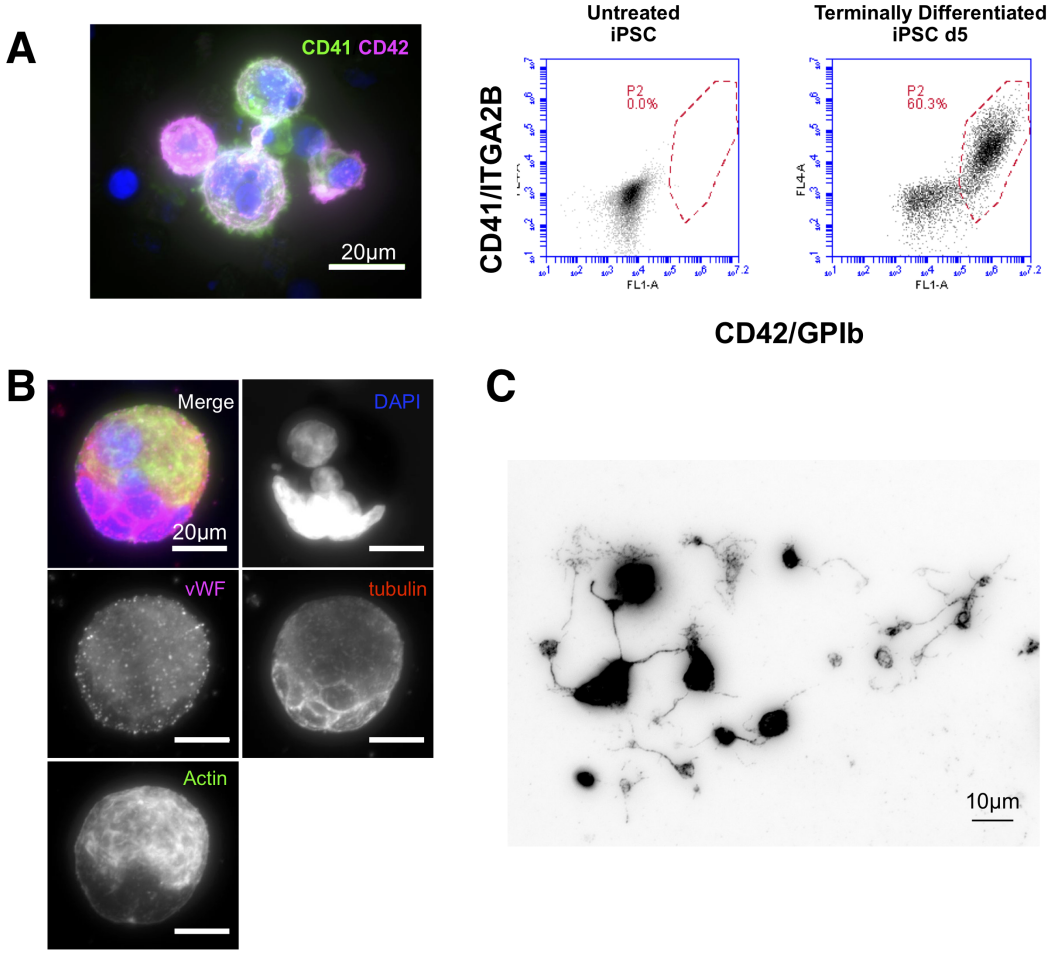

Figure 3.15: **Initial validation of iPSC MK cultures.** Initial validation of iPSC MKs through FACS and immunfluorescence confirms characteristic hall marks of megakaryocytes. (A) CD41/CD42 staining, the archetypal markers of mature MKs, reveals a substantial population of iPSC MKs (60%) at day 5 of the terminal differentiation. When spun down and imaged on Poly-L-Lysine, round healthy cells double positive for both markers are observed. (B) Immunofluorescence further reveals the presence of large, lobed nuclei. These polyploid cells are typical of megakaryocytes. vWF, a platelet granule protein, can also be seen accumulating in the cell. (C) Finally immunofluorescence of platelet extensions in this tubulin stained image shows a fine network of beaded proplatelet extensions.

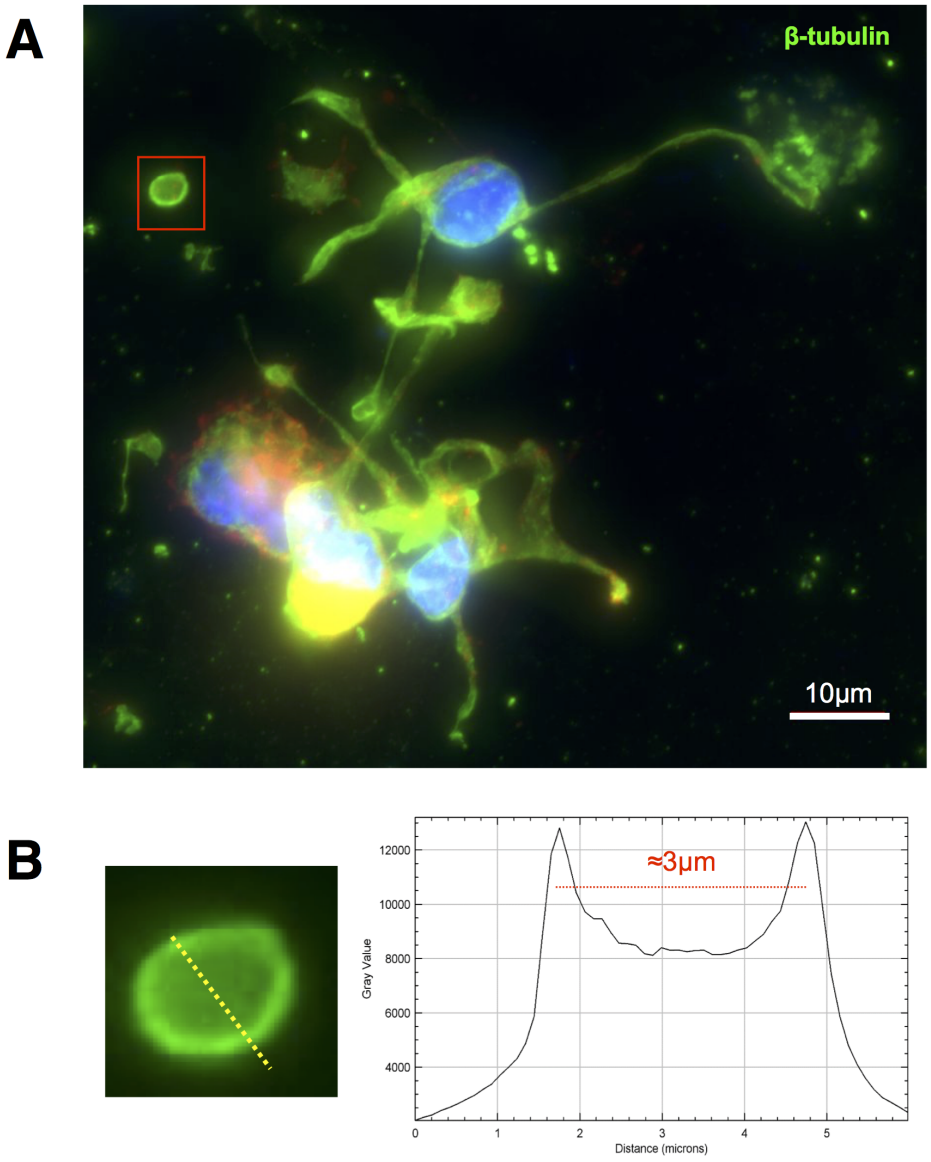

Figure 3.16: **Platelet production in iPSC-MKs.** (A) Platelet like particles are observed in cultures as iPSC MKs mature and develop proplatelet extensions. (B) These cells are discoid with a distinctive marginal band, and appear approximately 3μm in size.

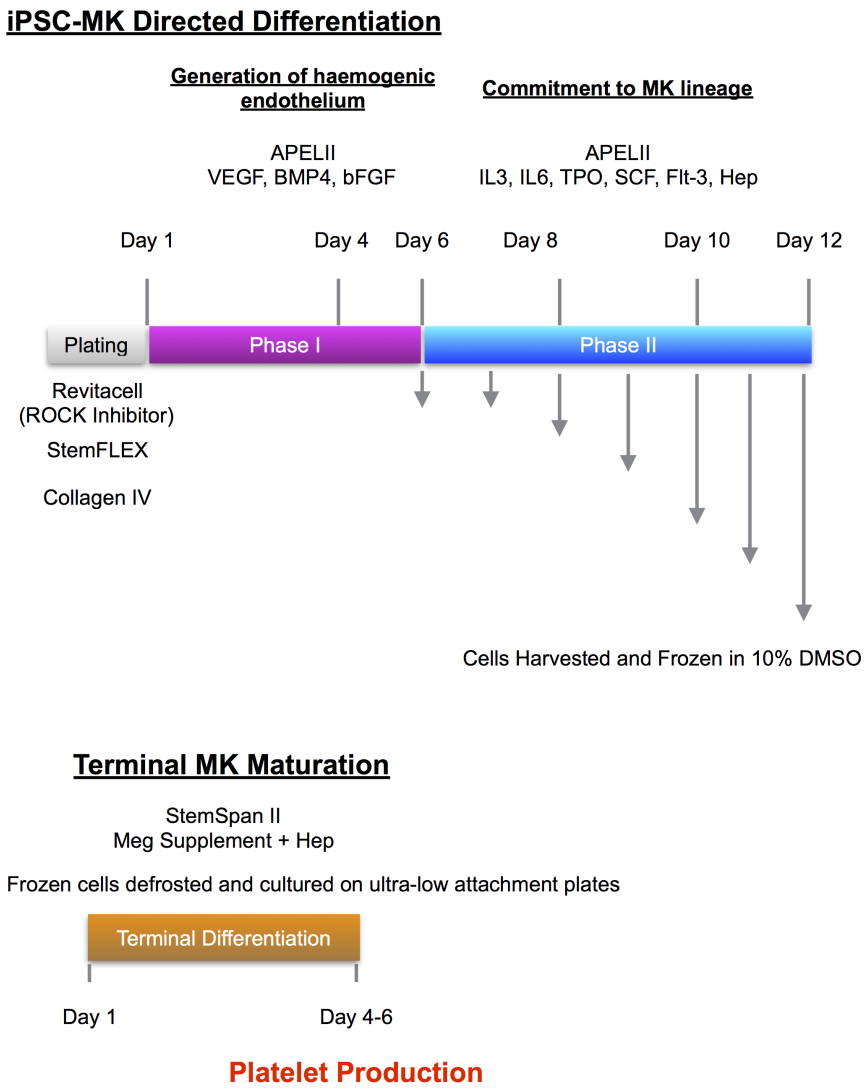

Figure 3.17: **Summary of revised directed differentiation protocol.**] The refined iPSC MK differentiation protocol developed in this chapter is summarised as a three phase process to yield mature, proplatelet forming cells. First, iPSC are seeded on Collagen IV with ROCK inhibitor in StemFlex medium. Cells are then placed in phase I medium comprised of APELII with VEGF, BMP4, bFGF (at 50ng/mL). Phase II involves altering the cytokine culture to favour haematopoeitic differentiation. This cocktail includes IL-3 and IL-6 at 10ng/mL, heparin at 5U/mL, and TPO, SCF, and Flt-3 at 25ng/mL. Progenitor cells are harvested and frozen in 10% DMSO on each day of phase II, before defrosting for the final, terminal differentiation. At this stage cells are cultured in StemSpan II with a megakaryocyte expansion supplement for 5-6 days, at which point platelet production can be observed in days 4-6.

3.5 Discussion and Conclusions

This chapter has reported on attempts to develop protocols for the CRISPR mediated editing of iPSC through HDR and for the differentiation of iPSC to mature MKs suitable for phenotypic screening. While much of the work targeting *TubA1B* has failed to yield a practical level of integration, the method of iPSC differentiation based on work by Feng *et al.* has produced a means by which to study platelet production *in vitro*.

3.5.1 Adapting CRISPR methods to iPSC

At the start of the project the literature primarily suggested that methods used to target cell lines could be conveniently transferred to iPSC - albeit at significantly lower efficiencies due to the poor transfection of these cells and other technical challenges[246]. In our work we have found that the relatively high level of knock-in we observe at the *TubA1B* locus in cell lines is not effectively reproduced in iPSC, even after adapting our methods to produce high transfection rates of RNPs reported to significantly increase editing efficiency.

While the use of Lipofectamine STEM and RNP complexes allows for high transfection efficiency, this has not translated to an improved efficiency of knock-in. The use of SCR-7, a small molecule inhibitor of NHEJ, only marginally improved HDR efficiency, while the use of CRISPY unfortunately had no effect on targeting at this site.

A notable paradigm shift is evident in more recent research, with more recent publications describing the need to optimise knock-in for specific cell systems. The CRISPY system is an excellent example of this, where the authors detail a method which substantially enhances HDR in iPSC, while having no effect on cell lines at all. This is supported to an extent in our work where we see a substantial increase in HDR in Hek293T on SCR7 treatment, but do not observe this in iPSC under the same conditions.

There are notable differences in our system which may account for the lack of improvement in HDR observed. The authors use an inducible Cas9 nickase system which ameliorates the

need for transfection, and will result in the expression of large amounts of Cas9 protein. As we achieve high rates of transfection, this is unlikely to exclusively account for the poor effects observed. The authors use a large single stranded oligonucleotide donor rather than double stranded DNA or a plasmid donor[260]. There is some evidence to suggest that editing efficiency is dramatically improved where single stranded DNA is used, and current efforts are focussed on generating ssDNA templates to trial insertion efficiencies.

Finally gene target is a key consideration when evaluating editing efficiency. In these optimisations we specifically target *TubA1B*, however it has been well established that editing efficiency is also dependent on the availability of target DNA in particular cell types[261]. Work in chapter 4 describes efforts to target the *TubB1* gene, but to effectively develop iPSC specific methods of gene tagging we would need to trial various editing strategies across multiple gene targets.

3.5.2 Adapting methods of iPSC differentiation to MKs

Substantial advances have been made in the development of protocols to generate mature megakaryocytes from iPSC. Both forward programming and directed differentiation approaches have advanced significantly, driven largely by the potential for transfusable clinical grade platelets. However, key limitations remain the relatively minimal production of platelets from these cells, particularly in the absence of a bioreactor. In terms of using iPSC-MKs for disease modelling, the lack of platelet production in these cells corresponds to poor proplatelet formation which limits the applicability of these cells for phenotypic screening.

Our initial attempts at reproducing the Feng *et al.* protocol would yield granular cells which, while positive for CD41 and CD42, would not produce proplatelets or become significantly larger over the course of a terminal differentiation (data not shown). Experimenting with different media quickly showed that APELII sustains substantially healthier cells which rapidly grow larger and form proplatelets in culture at the end of the terminal differentiation

phase. Feng *et al.* reported this, but do not show quality imaging to support the method as a means of generating a large sample of proplatelet forming cells.

Our CD41/CD42 double positive fraction appears larger than that which is reported by Feng *et al.* (60.3% compared to 37%), though this may be a consequence of preferential differentiation towards this lineage in our specific iPSC line. Epigenetic memory is known to influence the efficiency of terminal differentiations in iPSC and is dependent on their tissue of origin[262].

Significant work remains to be done to finalise this protocol. Further FACS and immunofluorescence to quantify the presence of other key MK markers will be needed. Similarly we would need to establish which portion of the population goes into platelet production to better optimise this approach as a phenotypic screen. However the immunofluorescence images generated thus far surpass those presented by other work demonstrating iPSC-MKs *in vitro* thus far, setting a solid foundation for the application of this method for future research in disease modelling.

Chapter 4

Post-translational modification of β-1 tubulin isoform

In previous chapters the development of CRISPR, single molecule imaging, and iPSC differentiation methods have been discussed. The ultimate goal of this project is the application of these techniques to the study of inherited thrombocytopenias. In this chapter these techniques are applied to the study of *TubB1* mutations identified in patients from the GAPP study.

4.1 Introduction: *TubB1* and the tubulin code

Microtubules are a key component of the cytoskeleton, involved in an array of functions including cell division, cargo transport, motility, and the development of specialist cell morphologies and functions[263, 264].

The building blocks of microtubules are heterodimers comprised of α- and β-tubulin monomers, each with approximately 50% sequence homology at the amino acid level and with a mass of approximately 50kDa [113]. α- and β-tubulin subunits polymerise to create linear arrays called protofilaments (Figure 4.1), 13 of which associate laterally to form cylindrical pseudo-helical, hollow microtubules 25nm in diameter[113].

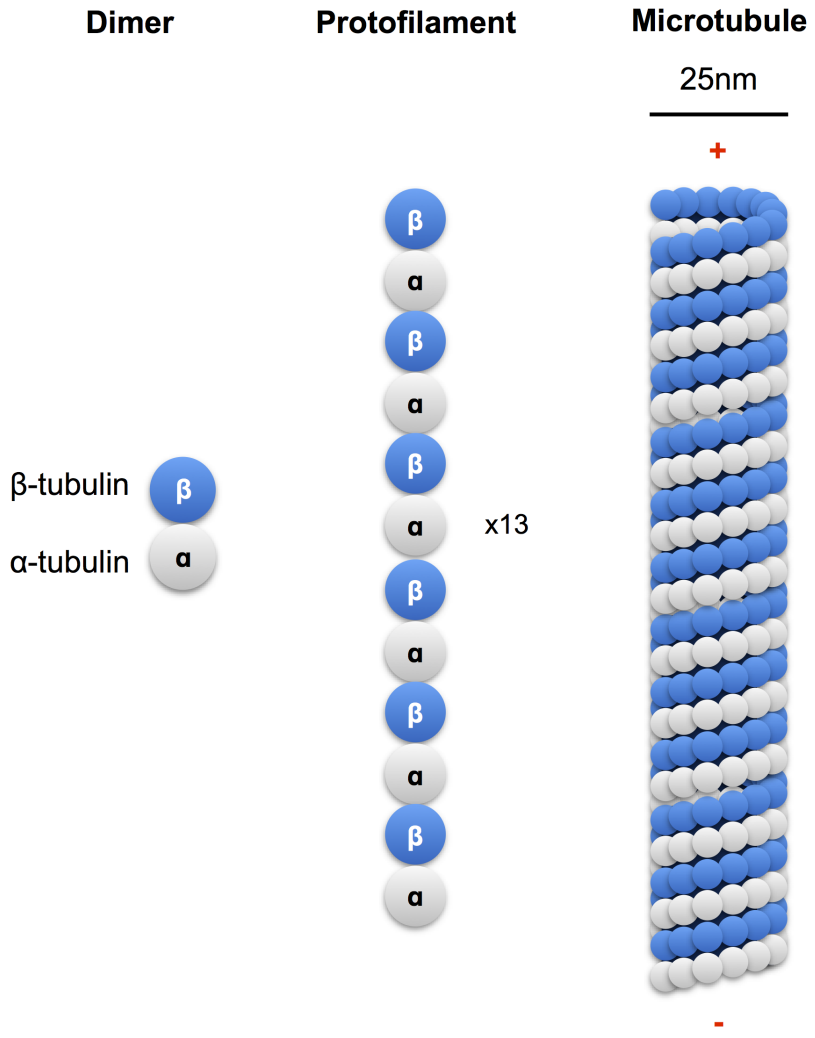

Figure 4.1: **Microtubule structure.** The building blocks of all microtubules are α- and β-monomers which assemble to form heterodimers driving the polarity and orientation of MTs. Each dimer is assembled into a linear array known as a protofilament, 13 of which associate laterally to form a stagged, pseudo-helical, hollow cylinder known as a microtubule. Each microtubule is 25 nm in diameter, and features a (-) and (+) end capped with an α- and β subunit respectively.

Polarity is key to microtubule function, with (-) and (+) ends terminating in α- and β-subunits respectively. Microtubules elongate in both directions, but polymerise more rapidly in the (+) direction. Polarity is a critical aspect of a vital property of microtubules known as 'dynamic instability' - namely the ability to polymerise and depolymerise extremely rapidly to achieve growth and shrinkage[263, 264].

When polymerising, tubulin dimers are GTP bound. While the GTP bound to α-tubulin is stable, the β-tubulin bound GTP can be hydrolysed after assembly - the subsequently formed GDP bound β-tubulin is susceptible to depolymerisation. Only GDP-bound β-tubulin subunits at the exposed end of a microtubule can dissociate, thus achieving rapid depolymerisation from the (+) end of the filament. Conversely, tubulin dimers assemble on to GTP-bound subunits, and so regulation of the GTP/GDP state of the ends of microtubules is key to their ability to rapidly grow and shrink[263, 264].

Microtubules are nucleated and assembled at organising sites known as MTOCs (Microtubule Organising Centres) which are comprised of a distinctive γ-tubulin[265]. γ-tubulin complexes with a host of accessory proteins to form a γ-tubulin ring complex (γ-TuRC) which acts as an organising and initiating template for dimer polymerisation in the (+) end away from the MTOC[265].

4.1.1 Post-translational modification of microtubules

As ubiquitous cellular structures, the question of how microtubules are regulated to achieve highly specialised functions and morphologies in specific cell types is a complex one. In recent years this question has been addressed by "the tubulin code", a paradigm which describes how MTs achieve unique roles through the differential expression of specific isoforms of α- and β-tubulin, and the subsequent post-translational modifications of these isoforms to regulate MT interactions with motor proteins, and other microtubule accessory and microtubule interacting proteins (MAPs and MIPs respectively)[113, 263, 264].

The expression of specific tubulin isoforms is a key driver of specific functions and morphologies[113]. Tubulin sequences are highly conserved across species, with abnormalities poorly tolerated and with significant functional consequences[266]. As such, differences between particular isoforms has significant consequences on microtubule dynamics, assembly, and ultimately function[113]. Variations in flexible regions like the C-terminal tail result in a mosaic of post-translational modifications which remain poorly understood. There are nine mammalian isoforms for each tubulin monomer[113]. Disease modelling has been a significant driver in understanding the effects of tubulin, and this is discussed in the context of post-translational modifications in this section. A summary of PTMs and their known role in human disease can be found in table 4.1.

Tyrosination and de-tyrosination refers to the alternating addition (ATP-dependent and TTL (tubulin tyrosine ligase) mediated) or removal of C-terminal tyrosine residue from α-tubulin (and the subsequent exposure of a terminal glutamate residue), with significant consequences on the stability of MTs[267]. Following detyrosination further amino acids may be removed to generate δ2-tubulin or δ3-tubulin, leaving the altered residue unable to undergo further tyrosination or detyrosination[268, 269].

Detyrosination of MTs is typical of stable microtubules, but is thought to be a consequence of stability rather than a cause[270]. Detyrosinated microtubules are typically more stable as certain kinesin motor proteins which mediate depolymerisation (e.g. MCAK (Mitotic centromere-associated kinesin (MCAK) and kinesin family member 2A (KIF2a)) preferentially target tyrosinated residues[271]. In contrast, KIF5 preferentially binds detyrosinated microtubules in a manner thought to facilitate a specific cellular transport role[272].

The interaction between detyrosinated microtubules and kinesin motors is vital to neuronal development, with evidence suggesting that differential populations of tyrosinated and detyrosinated MTs in axons drive specific MT dynamics and cargo transport in axons and growth cones[113, 263, 273] The loss of tyrosination evidenced by TTL knock out mice results in

PTM	Sub-types of MTs modified	Reported defects and links with human disease
Detyrosination	Centrosomes, mitotic and meiotic spindles, axonemal MTs of cilia and flagella, all neuronal microtubules except the growth cone, marginal band of platelets, most myocyte microtubules	Dysfunction in tyrosination associated with poor cancer prognosis (neuroblastoma and breast cancer in particular). Loss of PTM results in neuronal abnormalities in mouse development. In myocytes role for mechanotransduction, and involved in cardiomyocyte contraction facilitates load bearing. Loss of the regulation of this PTM can result in arrythmia and muscle injury in Duchenne Muscular Dystrophy.
Acetylation	Mid-body and spindle microtubules, enriched on axonemal MTs of cilia and flagella, found in all neuronal microtubules except growth cone. Located on platelet marginal band and found in proplatelet extensions of megakaryocytes.	Loss of this PTM causes subfertility in mice, and defects in neurological development. Loss of acetylation corresponds with a loss of axonal development in Charcot Marie Tooth disease. Loss of acetylation causes thrombocytosis in patients with myeloproliferative neoplasms.
Polyglutamylation	Spindle and mid-body microtubules, enriched on centrioles. Abundant in axonemal MTs (cilia and flagella), enriched on neuronal development.	Loss of polyglutamylation is linked to loss of male fertility as a consequence of poor spermatogenesis and sperm motility. Polyglutamylation is enriched in photoreceptors, loss of this polymodification has been linked to retinopathies and other visual disorders. Affects respiratory cilia in mice. Loss of this PTM causes extensive neurodegeneration in mice.
Polyglycylation	Enriched on primary cilia and motile cilia. Poorly studied if at all in other cell types.	Implicated in colon cancers, loss results in retinal degeneration in mice.

Table 4.1: **Summary of tubulin PTMS and their roles in human disease.** While the tubulin code remains a relatively recent field, the role of PTMs in healthy cell function and development has been highlighted by human disease models. The mechanism by which many of these PTMs regulate cell function remain unknown. Adapted from Magueira *et al.* 2018. Found in myocytes though function remains unclear. While other tubulin PTMs have been reported, they are not as strongly linked to human disease.

significant defects in brain architecture and peri-natal death, likely due to defective neurite outgrowth[274].

Tyrosination also has a role in muscle cells, where detyrosination has been shown to affect the function of MTs in cardiomyocyte contraction[275]. A loss of detyrosination results in cardiac dysfunction by altering the stiffness of cardiomyocytes, while the converse upregulation of detyrosination is found in patients with cardiomyopathy[276].

Acetylation is a relatively well understood PTM which typically targets lysine 40 (K40) of α-tubulin monomers, which lies within the lumen of MTs and is mediated by αTAT1, a tubulin acetyle transferase[277]. Acetylation of K40 is reversed by either sirtuin2 (SIRT2) or histone deacetylase 6 (HDAC6)[278]. Other acetylation sites have been more recently characterised, however their function has yet to be understood[279].

Neuronal MTs are heavily acetylated to mediate axon branching in development and the migration of cortical neurons[273]. A recent report by Xu *et al.* shows that increased acetylation results in an increased resistance to mechanical stress. Decreased tubulin acetylation has other significant neurological consequences, for example decreased acetylation has been linked to decreased axonal transportation in a range of diseases including Huntington's Chorea, Charcot-Marie-Tooth, and Parkinson's disease[280, 281, 282].

Polyglutamylation and glycylation are PTMs which target glutamate residues on both tubulin monomers and result in the addition of glutamate or glycine residues respectively[113, 263]. Interestingly polyglutamylation has been observed in microtubules in centrioles, axenomes, neuronal outgrowths, and mitotic spindles[113, 263].

Thus far polyglycylation has primarily been observed in axenomes, thus suggesting a role for polyglycylation and polyglutamylation in regulating ciliary function - with important consequences for ciliopathies[113].

As these polymodifications target the same specific substrate - namely glutamate residues in tubulin tails, there has been a level of competition observed between these PTMs. For

example, glutamylation is evident on β-tubulin in post-natal development, but is found on α-tubulin in younger neurons[283]. There is some debate as to whether these polymodifications negatively regulate one another[284, 285].

Ciliopathies result in a range of disorders which stem from aberrant functions in both motile cilia and flagella, but also primary cilia[286]. The polyglutamaylation of C-terminal tails has been shown to be an important regulator of ciliary beating[273], with mutations in the glutamylases and deglutamylases regulating this polymodification causing male infertility through aberrant spermatogenesis and poor sperm motility[287, 288, 289]. Other motile cilia include airway cilia , whose function is impaired on the loss of the polyglutamylase TTLL1 in mice[290].

Primary (or non-motile) cilia are vital sensory organelles with important signalling roles. In mice with a variant of the deglutamylase CCP1, an increase in polyglutamylation results in the degeneration of photoreceptors in the retina[291]. A series of human mutations in the deglutamylase CCP5 have been reported in patients with visual impairments[292, 293].

A loss of glycylation has similarly been reported to affect ciliary function and length in mice, and a loss of this PTM in photoreceptors results in ciliary shortening and subsequent retinal degeneration in mouse models[294, 295]. A knock-out model of TTLL3 results in a loss of glycylation and the development of tumours in the colon[296].

Beyond ciliopathies glutamylation has significant roles in neuronal development and function. Current research suggests a role for polyglutamylation in axonal transport (due to the effects of this PTM on motor protein processivity) and synaptic function[297].

Amongst the major modifications discussed above, there are a host of other PTMs which have been discovered but remain poorly reported. These include phosphorylation, polyamination, glycosylation, sumoylation, arginylation, ubiquitylation and methylation[298].

4.1.2 Motor proteins

Microtubule function is highly dependent on the recruitment and subsequent actions of microtubule associated proteins (MAPs), of which motor proteins are a key component. Referred to briefly in the section above, the microtubule motors are (-) end driven dynein and (+) end driven kinesin[299].

Dyneins are split into two broad categories, axonemal and cytoplasmic dyneins[299]. Axonemal dyneins are so named for their role in force generation in cilia, and this class of motor protein is thought to drive microtubule sliding.

Cytoplasmic dyneins are further divided into two classes, cytoplasmic dynein-1 and cytoplasmic dynein-2. Cytoplasmic dynein 1 is expressed ubiquitously in mammalian cells and associated with various microtubule dependent processes, including vesicular transport, cell migration, nucleokinesis, and nuclear envelope breakdown amongst a wealth of other critical cell functions[299]. Cytoplasmic dynein 2 is involved in transport within specialised structures like cilia and flagella, which are in turn facilitated in their unique mechanical actions by a separate class of axonemal dyneins.

Kinesins are a primarily (+) end driven motor protein with key, ATP mediated, roles in cell division and cargo transport[300]. Prominent examples include the role of of kinesin sub-types in transporting neurotransmitters along the length of axons. Recently kinesin-1 has been reported as a player in platelet secretion[300].

4.1.3 Microtubule accessory proteins

Microtubule accessory proteins, usually abbreviated to MAPs, are known to be regulators of microtubule dynamics[301]. Tau proteins are MAPs with weights below 55-62 kDA and are known to prevent depolymerisation and promote nucleation. In axons tau proteins have been implicated in Alzheimers, with a role in stabilizing axonal MTs[302].

There are four MAPs with molecular weights between 200-1000 kDa (creatively named

MAP-1, MAP-2, MAP-3 and MAP-4). MAP-1 C is actually cytoplasmic Dynein, MAP-4 is found in most mammalian cells, and MAP-2 is specific to the body of dendrites where they mediate interactions with other cytoskeletal filaments[301]. Katanin, spastin, and figetin are known as destabilisers of MTs (through microtubule severing) and therefore regulate tubule length[301].

Plus end tracking proteins are MAPs which find to growing MTs and are involved in MT interactions with chromsomes. TIPs include End Binding proteins (EB-1, EB2, EB3), Dynamitin, Lis1, and CLASP1 and 2[301].

4.1.4 Microtubules in platelets and inherited thrombocytopenia

Critical to achieving these functions is the expression of *TubB1*, the platelet and megakaryocyte specific isoform of β-tubulin. As detailed in section 4.1, microtubules play a critical role in proplatelet formation, and hence platelet production. Knock-out mouse models and a series of studies highlight the importance of *TubB1* in achieving the unique cytoskeletal changes needed for MKs to produce platelets.

TubB1 is located on chromosome 20 and is 7,401 base pairs in length[303]. The resulting protein, tubulin β-1 is 451 amino acids in length, with a mass of approximately 50kDa[303]. Interestingly tubulin β-1 is the most divergent at the sequence level of the many mammalian tubulin isoforms, and features the longst C-terminal tail.

Platelets rely on microtubules to maintain a marginal band when quiescent, resulting in the classic flat discoid shape of circulating resting platelets[304, 305]. While other haematopoeitic cell types assemble marginal bands, platelets are the only known mature cell in humans which require a circumferential marginal band[306]. The MB is comprised of 7-12 rings of mixed polarities actively assembled during both resting and activation[307]

The activation of platelets involves a tightly regulated rearrangement of the cytoskeleton which results in a series of shape changes. Antagonistic motor proteins maintain the resting

state of the marginal band, and during platelet activation a motor protein dependent mechanism results in sliding which extends the marginal band and causes the transition to a spherical shape [305]. This change can either be transient and reversible or, in the presence of a sufficiently strong stimulus, become irreversible through actomyosin contraction and the subsequent concentration of granules to the centre of the platelets[308]. The disc-to-sphere transition is a critical part of platelet activation, which also involves the secretion of granules as a terminal step in platelet activation.

In humans, *TubB1* mutations have been shown to result in impaired platelet production, with a resulting macrothrombocytopenia (a list of reported rare variants and their phenotypes is found in table 4.2) An early report of a *TubB1* mutation resulting in a familial congenital macrothrombocytopia by Kunisima *et al.* described low platelet counts (40-60 x 10^9/L) in

Variant	Protein	AF	Phenotype	Ref
c.35delG	p.Cys12Leufs*12	0.0025	Macrothrombocytopenia	Bastida *et al.* 2017
c.319A>C	p.Thr107Pro	0.0016	Macrothrombocytopenia	Bastida *et al.* 2017
c.326G>A	p.Gly109Glu	0.0873	Macrothrombocytopenia	Auer *et al.* 2014, Astle *et al.* 2016
c.721C>T	p.Arg241Trp	0.0107	Thrombocytopenia (Platelet size n/a)	Johnson *et al.* 2016
c.726C>G	p.Arg242Leu	-	Thrombocytopenia (Platelet size n/a)	Johnson *et al.* 2016
c.779T>C	p.Phe260Ser	-	Macrothrombocytopenia	Kunishima *et al.* 2014
c.860C>T	p.Ala287Val	0.0058	Increase in platelet distribution width (GWAS)	Astle *et al. 2016*
c.952C>T	p.Arg318Trp	0.0041	Macrothrombocytopenia	Kunishima *et al.* 2009
c.1075C>T	p.Arg359Trp	0.6233	Macrothrombocytopenia	Bastida *et al.* 2017
c.1080insG	p.Leu361Alafs*19	-	Thrombocytopenia (Platelet size n/a)	Johnson *et al.* 2016
c.1267C>T	p.Gln423*	0.0017	Macrothrombocytopenia	Bastida *et al.* 2017, Fiore *et al.* 2017

Table 4.2: **Summary of reported *TubB1* mutations.** A summary of reported rare *TubB1* variants. A range of mutations across the length of the protein typically result in macrothrombocytopenia. Table adapted from Burley *et al.* 2017.

a 7 year old patient initially diagnosed with idiopathic thrombocytopenia purpurra. This novel p.R318W mutation resulted in abnormally large platelets, however platelet aggregation, megakaryocyte number and function also appeared normal. The same group followed up with a separate report in 2014 of a p.F260S mutation with a similar phenotype [309]. More recently, a C-terminal truncation of β-tubulin 1 has been shown to cause a macrothrombocytopenia, suggesting that C-terminal modifications may be drivers of protein function and causative of the disease phenotype observed [310].

Genome wide association studies (GWAS) have shown that *TubB1* variants can be broadly linked with changes in mean platelet volume and distribution width correlating with macrothrombocytopenias reported in the GAPP cohort and other studies[311, 312]. Interestingly the majority of GWAS *TubB1* variants can be found in non-coding regulatory elements.

While the link between *TubB1* variants and macrothrombocytopenia has been established, the question of how the loss of protein function results in defective proplatelet formation has yet to be addressed. Similarly the question of whether or not there is a functional effect on platelet function has not been fully addressed. Fiore *et al.* show that the C-terminal is likely to be a vital facilitator of MT function in platelet production, but like the majority of studies investigating *TubB1* variants the limited availability of patient sample and expression of mutated protein through over-expression in cell lines severely restricts any mechanistic interrogation of interesting variants[310].

4.1.5 Aims and Hypotheses

In the context of the tubulin code, *TubB1* is an archetypal example of a lineage restricted isoform, the expression of which results in unique morphologies and functions executed by motor proteins. Despite this, and a strong link to human disease, the nature of *TubB1* specific PTMs remains unknown, and while some elegant work has been performed to highlight the role of motors in executing both platelet production and function on activation, the link

between *TubB1* and these consequences remains relatively poorly understood.

Therefore the aim of this chapter is to apply methodologies developed thus far over the course of this studentship to establish the role of C-terminal PTMs in proplatelet production and subsequently in the activation and spreading of platelets. Thus far the only reported PTMs in platelet MTs are acetylation and tyrosination, neither of which are specific to *TubB1* and are therefore, while undeniably important, unlikely to be effectors of MT function in these cells.

Using patient mutations as a model, we aim to interrogate likely *TubB1* specific modifications at both the proplatelet and platelet spreading stage. The C-terminus of β-tubulin is rich in glutamate residues (figure 4.2), and therefore is a likely site for polyglutamylation and glycylation - neither of which have been reported in platelet or megakaryocyte PTMs.

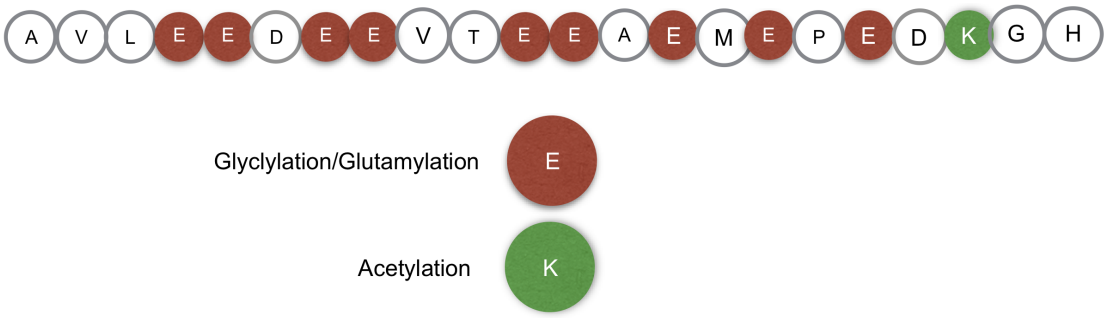

Figure 4.2: **Sequence of the β-tubulin 1 C-terminus.** The unique C-terminus of *TubB1* is rich in glutamates making it a likely site for glutamylation and glycylation, neither of which have been reported in megakaryocytes or platelets. We reason that these modifications are likely to be important to both platelet production and function by co-ordinating motor proteins and their associated sliding and cargo transport roles.

4.2 Materials and Methods

4.2.1 Platelet preparation

Whole blood was obtained for each experiment from healthy volunteers under the University of Birmingham's ERN 11-0175 license 'The regulation of activation of platelets'. Volumes of 25 mL were drawn from volunteers into sodium citrate. PRP (platelet-rich plasma) was generated by centrifugation of samples for 20 minutes at 200g. PRP was further spun to isolate platelets by centrifugation at 1,000g for 10 minutes with prostacyclin (0.1 μg/mL) and ACD. The resulting pellet was suspended in Tyrode's buffer prepared fresh, and pre-warmed at 37^0C (5mM glucose, 1mM MgCl2, 20mM HEPES, 12mM NaHCO3, 2.9 mM KCL, 0.34mM Na2HpO4, 129mM NaCL to a pH of 7.3). This suspension was spun again at 1,000g with prostacyclin at the same concentration before re-suspended to a final concentration of 2 x 10^8. Platelets were left to rest for 30 minutes at room temperature before any further processing or treatment.

The UK-GAPP study is included in the National Institute of Health Research Non-Malignant Haematology Study Portfolio (ID-9858) and was approved by the National Research Ethics Service Committee of West Midlands (Edgbaston REC reference: 06/MRE07/36). Participants gave written informed consent in accordance with the Declaration of Helsinki. The study was registered at www.isrctn.org as ISRCTN 77951167.

4.2.2 Platelet spreading

Resting platelets were fixed by preparing platelets at a concentration of 4 x 10^7 and mixing with equal volumes of 10% neutral buffered Formalin in a 15mL falcon tube. This mixture was inverted gently to mix the sample, and left to incubate for 5 minutes before subsequently adding 300μL of the resulting fixed, resting platelets to coverslips coated in Poly-L-Lysine (Sigma). Cells were then spun down at 200g for 10 minutes.

Spreading was performed on human fibrinogen (Plasminogen, von Willebrand Factor and Fibronectin depleted - Enzyme Research Laboratories) and Horm collagen (Takeda). Coverslips were coated overnight at a concentration of 100 μg and 10μg/mL (fibrinogen and collagen) respectively, before blocking for 1 hour in denatured fatty acid free 1% BSA (Life Technologies). Finally coverslips were washed once with PBS before the addition of platelets. Unless otherwise stated (as in the time course experiment), platelets were spread for 45 minutes at 37^0C. Fixation for spread platelets was performed in formalin as for resting platelets for 10 minutes.

4.2.3 Immunofluorescence

After fixation platelets were washed twice with PBS before incubation in 0.1% Triton X-100 for 5 minutes. The subsequently permeabilised cells were washed twice with PBS before blocking in 2% Goat serum (Life Technologies) and 1% BSA (Sigma).

Fixed, permeablised, and blocked cells were then incubated with primary antibodies at a concentration of 1:500 unless otherwise stated. The following antibodies were used for experiments in this chapter: polyglutamylated tubulin (mouse monoclonal antibody, clone B3 T9822, Sigma), pan-polyglycylated antibody (mouse monoclonal antibody, AXO49, MABS276 Millipore), monoglycylated antibody (AXO 962 mouse monoclonal MABS277, EMD Millipore), kinesin-1 (rabbit monoclonal to KIF1B ab 167429, abcam), axonemal dynein, β-tubulin (Rabbit polyclonal PA5-16863 tyrosinated tubulin (rabbit monoclonal antibody, clone YL1/2, MAB1864, EMD Millipore), acetylated tubulin (Lys40, 6-11B-1, mouse monoclonal antibody, Cell Signalling Technology). DNAL1 antibody (PA5-30643 Invitrogen).

After a 1 hour incubation in the relevant mix of primary antibodies. Cells were washed with PBS 2x before incubation in secondary antibodies (Alexa568-phalloidin, anti-rabbit Alexa-647, anti-mouse Alexa-588) for one hour at a dilution of 1:300 in PBS.

4.2.4 Image and statistical analysis

Statistical analysis was performed using GraphPad PRISM 7. Image analysis was performed using a custom segmentation algorithm developed by Jeremy Pike. Briefly, the actin channel from resting and spread platelet images was used to train Ilastik classifiers (approximately 6 images per condition) for segmentation based on this channel. This was incorporated into a Knime workflow which would run images through the classifier to generated segmented binaries in which co-localisation and fluorescence intensity statistics were calculated. For the data presented in this chapter, M1 (a corrected Mander's co-efficient to channel 1) was used to determine the co-localisation of PTMs to tubulin, and an M2 value (corrected Mander's co-efficient to channel 2) was used to calculate the co-localisation of motor proteins to PTMs of interest.

4.2.5 Cloning

Donor vectors were generated as described in chapter 2. Briefly, homology arms and inserts were designed with 20 base pair overlaps and ordered as gBlocks from IDT. On arrival, gBlocks were resuspended and added at equimolar ratios to an empty pGem-T-Easy backbone for a 4 fragment gibson assembly (Hi-Fi kit, NEB). Clones were first validated by EcoRI digest before Sanger sequencing (UoB Genomics) to validate the correct assembly of each fragment.

Mutants of the *TubB1* overexpression plasmid were generated through the use of a Q5 site directed mutagenesis kit (NEB). For each reaction, five colonies were picked for validation by Sanger sequencing. All sequences for primers and homology arms for CRISPR experiments are described in 6.2.

4.2.6 iPSC MK generation

iPSC-MKs were generated as described in chapter 3. Immunofluorescence of iPSC MK samples was performed identically to platelet experiments.

4.3 Results

4.3.1 Identifying and modelling candidate *TubB1* mutations

Two C-terminal *TubB1* mutations were identified from the GAPP cohort, each of which belong to unrelated families presenting with macrothrombocytopenia (Figure 4.3). Family A patients were found to be heterozygous for a C to T missense mutation resulting in an arginine to tryptophan change (c.C1075T, p.R359W) in *TubB1.* Individuals in this family also carry a *GFI1B* mutation (p.Cys168Phe), however only individuals A:1 and A:3, both of whom carry both mutations, present with a macrothrombocytopenia (107 and 85 x 10^9/L). Individual A:2 has the GFI1B mutation but does not carry the TubB1 variant and presents with a normal platelet count (221 x 10^9/L). Interestingly, individuals A:1 and A:3 also present with significantly higher immature platelet fractions and mean platelet volumes than their *TubB1* WT relative (53.5% and 55.1% compared to 25.5%, MPV for A:1 and A:3 too large for measurement). This variation in count and phenotype suggest that *TubB1* is causative of the macrothrombocytopenia in this instance.

Family/patient B was an elderly gentleman (now deceased) with a G insertion and subsequent frameshift truncation of the protein 19 amino acids from the site of insertion (c. 1080insG, p.:361Afs*19). This patient had a severe macrothrombocytopenia with a count of 11 x 10^9/L and MPV of 13.4 (Table 4.3). At the time of study, IPF measurement was unavailable.

Both mutations are positioned in the C-terminal region of the $\beta - 1$ isoform encoded by *TubB1* as indicated in figure 4.3B. This region is positioned away from the dimer:dimer interface, and is a likely site for post-translational modification. Both affected nucleotides are highly conserved in mammals (Figure 4.3C). We predict that the R359W missense is likely to alter the fold of the C-terminal tail, potentially affecting PTM or interactions with critical MAPs (Figure 4.3 D). Similarly the G insertion and subsequent frameshift are likely

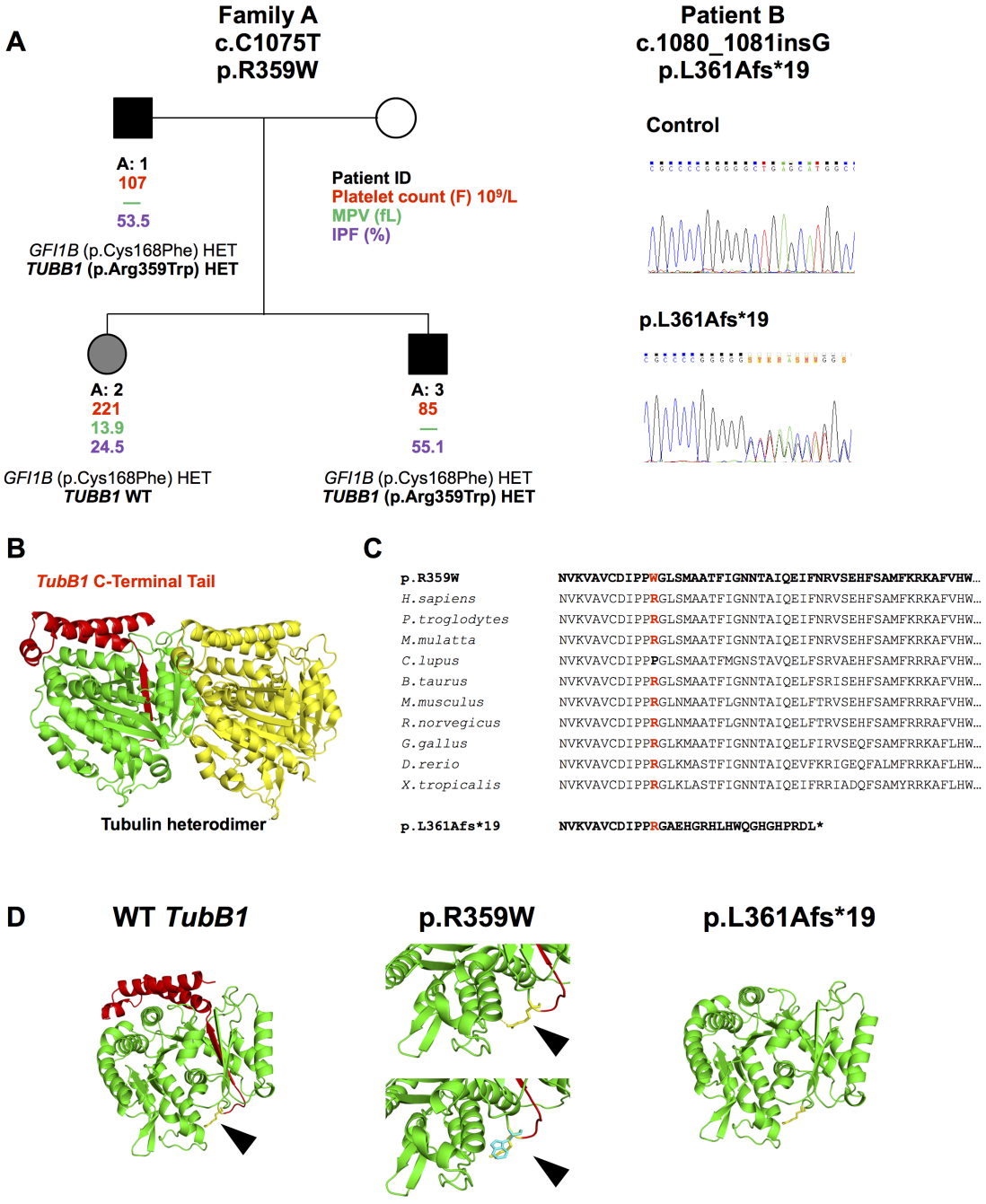

Figure 4.3: **Candidate *TubB1* mutations and their hypothesised effect on the C-terminus of β-tubulin.** (A) Two unrelated families were identified as carrying mutations in the *TubB1* gene within 6 base pairs of one another. The first, family A, is comprised of 3 individuals, two of whom carry an Arginine to Tryptophan (p.R359W) coding mutation.

Figure 4.3: *Cont.* Interestingly, all 3 individuals in family A harbour a *GFI1B* mutation. However, the individuals with the reported *TubB1* mutation (A:1 and A:3) present with a macrothrombocytopenia and high IPF, while the patient without the R359W *TubB1* mutation presented with a normal platelet count. The second family is comprised of a single individual, recently deceased, with a frameshift mutation 6 base pairs from the missense reported in family A. In this individual's case, the insertion of a guanine nucleotide results in a frameshift with a premature stop codon 19 amino acids from the leucine to alanine change. (B) The C-terminal tail is downstream of both mutations in these families, and projects away from the dimer/dimer interface. (C) The arginine residue mutated in family A is highly conserved across species, as are sequences adjacent to the frameshift in patient B. (D) Based on homology modelling of *TubB1*, we predict that the missense mutation reported in family A is likely to affect the fold of the C-terminal tail, while the frameshift causes a truncation of the C-terminal tail.

to truncate the C-terminal region if the stability of the protein is not undermined by such a significant deletion.

Patient	Age	*TubB1* mut.	Plat. Ct. (10^9)	MPV(fL)	IPF (%)	Secondary Defect
A:1 (M)	66	p.R359W	107	large	53.5%	Yes (Secretion)
A:2 (F)	41	-	221	13.9	24.5%	Yes (Secretion)
A:3 (M)	29	p.R359W	85	large	55.1%	Yes (Secretion)
B:1 (M)	dec.	p.L361Afs*19	11	13.4	N/A	Yes (Secretion, Gi)

Table 4.3: **Summary of Patient Data.** Family A is comprised of 2 affected individuals with the *TubB1* p.R359W mutation with macrothrombocytopenia (note an MPV of large refers to platelets too large to be accurately measured) and one unaffected individual with a normal platelet count. All individuals in this family were positive for a predicted non-disease causing *GFI1B* mutation (p.C168F). Interestingly the two affected individuals also have a much higher IPF than the unaffected individual in this family. Family B is comprised of a single individual recently deceased, with a severe macrothrombocytopenia, and both a secondary secretion and Gi defect.

The GAPP project includes FACS data as part of the initial phenotyping process, which provides an interesting functional insight into the effects of both mutations on the recruited patients. Interestingly, patient B shows a significant reduction in surface P-selectin and Fibrinogen uptake in response to all agonists (Figure 4.4). Family A shows no change in surface receptor expression, but show a weak P-selectin and fibrinogen response to low dose ADP, CRP, and PAR-1, suggesting a mild secretion defect (Figure 4.4 C,D).

To determine whether these mutations alter the fold of β-1 we first cloned the *TubB1* sequence into an mApple over-expression vector (Figure 4.5). After validation and sequencing, this vector was further mutated using a Q5 site directed mutageneis kit to incorporate a thrombin site (for cleavage on protein purification) and each individual patient mutation. Work to establish the CD of these proteins versus wild type is currently underway, and will establish whether the patient phenotype observed is an effect restricted to the C-terminal tail, or a consequence of a catastrophic misfold of the protein.

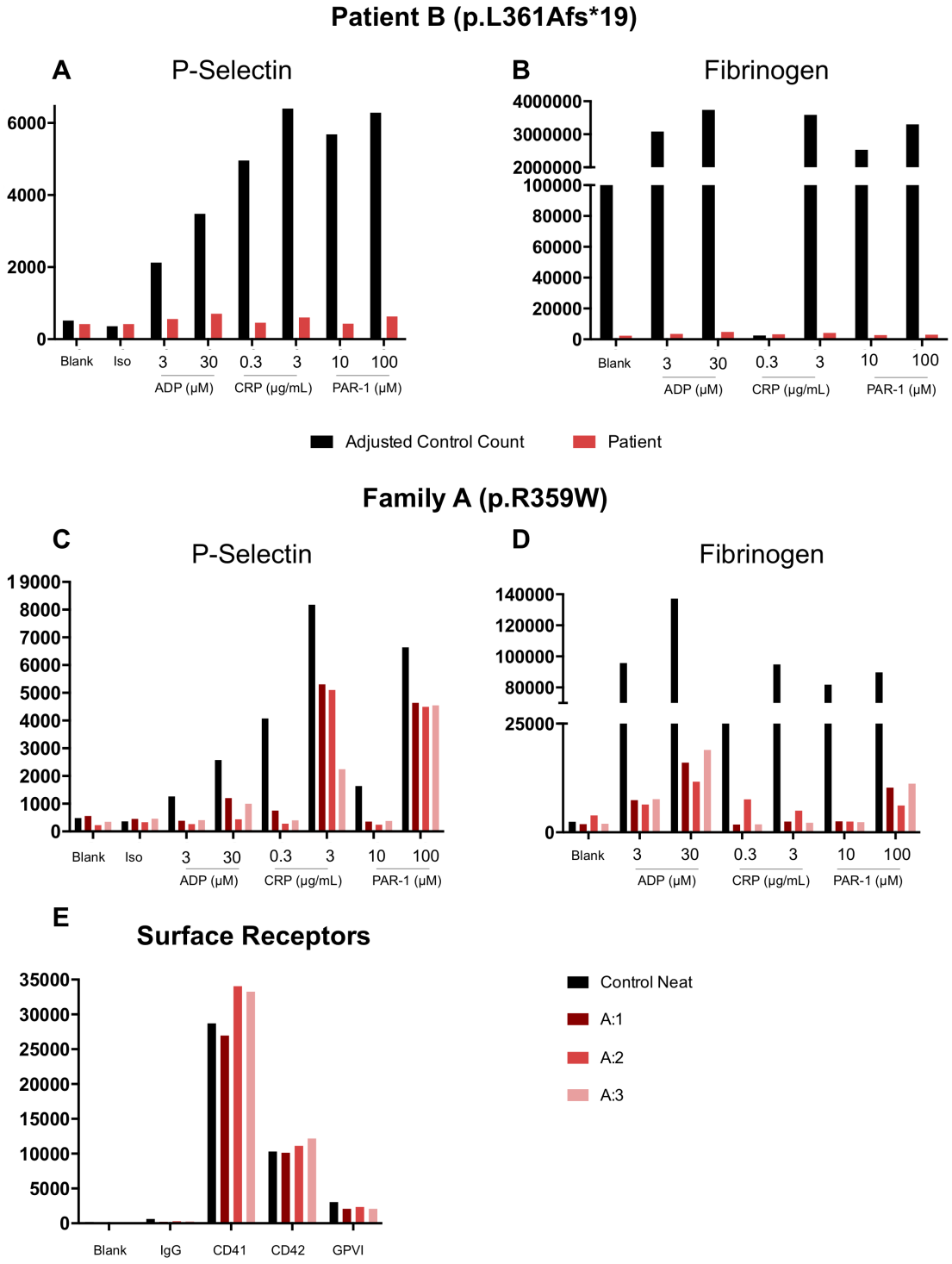

Figure 4.4: **Patient flow cytometry data reveals secondary defects.** The GAPP project collects phenotypic data on patient recruitment, allowing for the assessment of secondary defects through FACS screening.

Figure 4.4: *Cont.* (A) Patient B shows a marked reduction in P-selectin surface expression and (B) fibrinogen uptake compared to controls. (C) Individuals from family A show a reduction in P-selectin at both doses of CRP, low dose CRP, and low dose PAR-1. (D) Patients similarly show a reduction in fibrinogen uptake compared to controls, but show no difference in (E) surface marker expression. *Note that this data was procured by Ben Johnson and Dr. Gillian Lowe through the GAPP study. Control counts were normalised for patient B due to the patient's severely low platelet count.*

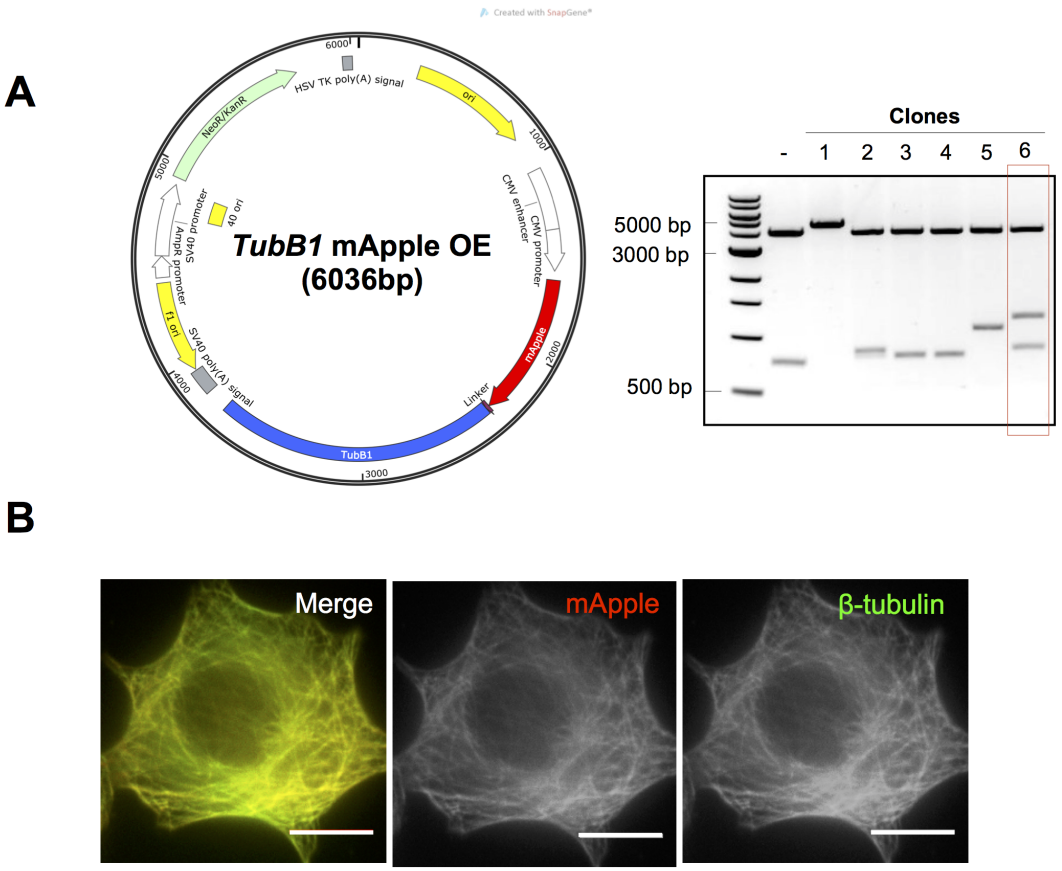

Figure 4.5: **Cloning of a fluorescently labelled β-1 tubulin over-expression vector.** (A) An N-terminally mApple tagged β-1 tubulin over-expression construct was designed to generate the protein for stability studies and for future transfection and analysis. The *TubB1* sequence was cloned into an C-terminal mApple empty vector generously provided by the Davidson lab. Upon gibson assembly, transformation, and the harvesting of individual clones, vectors were mini-prepped and analysed by EcoRI restriction digest. Clone 6 was verified has having correct cleavages at the predicted sites, and was sequenced to confirm a correct assembly. (B) The plasmid was then transfected and co-stained with a β-tubulin antibody to confirm its function. *10μm scale bar.*

4.3.2 Investigating candidate C-terminal PTMs on WT platelets

As we hypothesise that the C-terminal tail of *TubB1* is rich in glutamate residues ripe for polymodification, we first performed immunofluorescence assays on wild type platelets from healthy donors to establish whether glutamylation or glycylation are evident. To date, no evidence showing either of these polymodifications has been presented to our knowledge.

Platelets spread on fibrinogen and collagen were compared to resting platelets fixed and spun down to poly-L-lysine. Platelets in each condition were co-stained with Phalloidin-568, an antibody to $\beta - tubulin$, and an antibody to either polyglutamylated, monoglycylated, or polyglycylated tubulin (Figure 4.6).

Interestingly while resting platelets demonstrate low levels of each of these polymodifications, their distribution and intensity is dramatically altered on spreading (and hence, activation) (Figure 4.6). Microtubules on fibrinogen and collagen spread platelets appear to be heavily polyglutamylated, particularly at the marginal band of partially spread cells (Figure 4.6A). Monoglycylated and polyglycylated residues alter their distribution dramatically, becoming punctage and diffuse throughout the cell (Figure 4.6 B,C).

To quantify whether these polymodified residues, which do not appear to be part of the marginal band in resting platelets, are incorporated into β-tubulin on spreading activation quantification of a Mander's coefficient was performed across all three conditions. Unsurprisingly, low co-localisation is observed between each polymodification and β-tubulin in the resting state, and this significantly increases on platelet spreading for polyglutamylated (**** $p < 0.0001$, *** $p = 0.003$ between fibrinogen and collagen spread cells respectively) and monoglycylated residues (** $p = 0.0037$ in fibrinogen spread cells).

Interestingly this corresponds to a significant increase in intensity on spreading in polyglutamylated tubulin (** $p = 0.0070$, * $p = 0.0415$ fibrinogen and collagen spread respectively), consistent with the qualitative observation of extensive polyglutamylation of the marginal band and β-tubulin on spreading (Figure 4.6E). Interestingly the intensity of monoglycylated

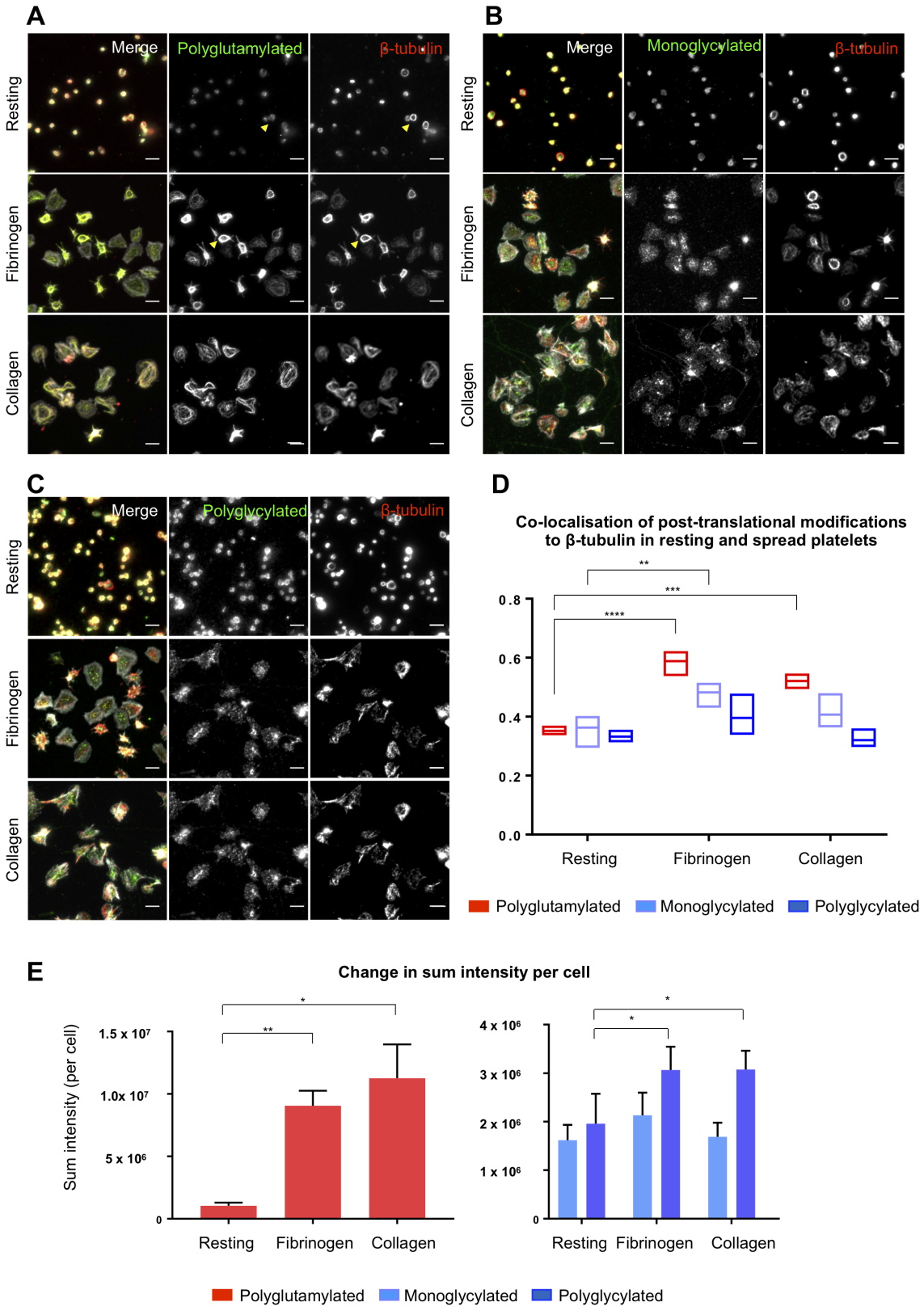

Figure 4.6: **Platelet spreading results in extensive post-translational modification.** (A) Resting platelets do not demonstrate a significant degree of polyglutamylation, particularly on the marginal band (yellow arrows).

Figure 4.6: *Cont.* On activation and spreading on fibrinogen and collagen, extensive polyglutamylation of the marginal band can be observed. (B,C) Both mono- and polyglycylated residues demonstrate a markedly different distribution on activation. (D) Quantification of the co-localisation of post-translationally modified residues and β-tubulin in both resting and spread platelets shows a statistically significant change across both post-translational modifications and conditions (** p = 0.0029 and **** p ¡ 0.0001 respectively). There is a significant increase in the co-localisation of polyglutamylated residues between resting and both fibrinogen and collagen spread cells (**** p ¡ 0.0001, *** p 0.003). A significant increase in the co-localisation of monoglycylated residues is also observed between resting and fibrinogen spread platelets (** p = 0.0037). (E) A significant increase in signal intensity per platelet is observed for fibrinogen and collagen spread cells versus resting control (** p = 0.0070 and * p =0.0415). No significant increase in monoglycylation is observed on platelet spreading, however interestingly a significant increase in polyglycylation is observed in both fibrinogen and collagen spread cells (* p = 0.0349 and * p = 0.0338). Two-way ANOVA with uncorrected multiple comparisons performed, line of each plot represents mean and min/max values. One-way ANOVA performed for polyglutamylated intensity measurements (S.D.) n = 3. *10μm scale bar.*

tubulin remains consistent on spreading despite an increase in colocalisation, while the extent of polyglycylation increases despite no significant change in co-localisation (* p = 0.0349 and * p = 0.0388 on fibrinogen and collagen spread cells respectively).

This data suggests that polyglutamylation, particularly of the marginal band, is a significant factor in the initial destabilisation of the tubulin ring required for rapid shape change on platelet activation. Interestingly we see an altered distribution of glycylated residues, and an increase in polyglycylation despite poor colocalisation to β-tubulin within the cell, suggesting that platelet spreading involves the extension of glycine chains on monoglycylated residues to potentially mediate MAPs.

To determine whether polyglutamylation precedes or eclipses other PTMs previously reported in platelets (namely acetylation and tyrosination), a time course was performed on fibrinogen to establish the temporal relationships between these modifications. Acetylation has been shown to decrease on platelet activation and spreading by Sadoul *et al.*. A rapid polyglutamylation of the marginal band and subsequent β-tubulin is evidence at every time point (Figure 4.7A), while a change in acetylation is more evident at later

time points (indicated by yellow arrows on the figure panel). Similarly no significant variation in tyrosination is observed. This trend is confirmed by the quantification of sum intensity per platelet across three replicates ($p < 0.0001$). This data suggests that extensive polyglutamylation of β-tubulin is indeed the driving PTM on platelet spreading.

The literature suggests that changes in motor protein function are likely drivers of marginal band deformation on platelet activation. To that end we investigated the effects of polyglutamylation on co-localisation with dynein, a motor known to be involved in both platelet production and function. As shown in Figures 4.7C and D, the polyglutamylation of tubulin coincides with a sharp decrease in co-localisation between dynein and polyglutamylated residues. This is consistent with evidence suggesting that hyperglutamylation has a significant effect on motor protein processivity.

Interestingly we note that dynein is present in leading edges as platelets spread, with this particular motor found at the periphery of fully spread cells as indicated by the yellow arrow in figure 4.7C. This distribution suggests that this (-) end driven protein is involved in driving the expansion of platelets on spreading.

Further co-localisation analysis was performed to determine the spatial relationship between dynein and other polymodified residues (Figure 4.8). On platelet spreading on fibrinogen and collagen, the same pattern of reduced co-localisation is observed in cells stained for polyglutamylated tubulin (** $p = 0.0082$ and *** 0.0004 in fibrinogen and collagen spread cells respectively). Surprisingly we do not observe this trend in glycylated residues, with a significant increase in co-localisation evident on fibrinogen spread platelets (* $p = 0.0251$). Polyglycylated residues do not show any change in co-localisation with dynein.

Kinesin is the second MT motor protein with critical roles in cargo transport. The reorganisation and subsequent secretion of platelet granules is a key aspect of platelet activation. Having observed a dramatic effect on the association of polyglutamylated residues with dynein, co-localisation studies were performed to investigate the spatial association

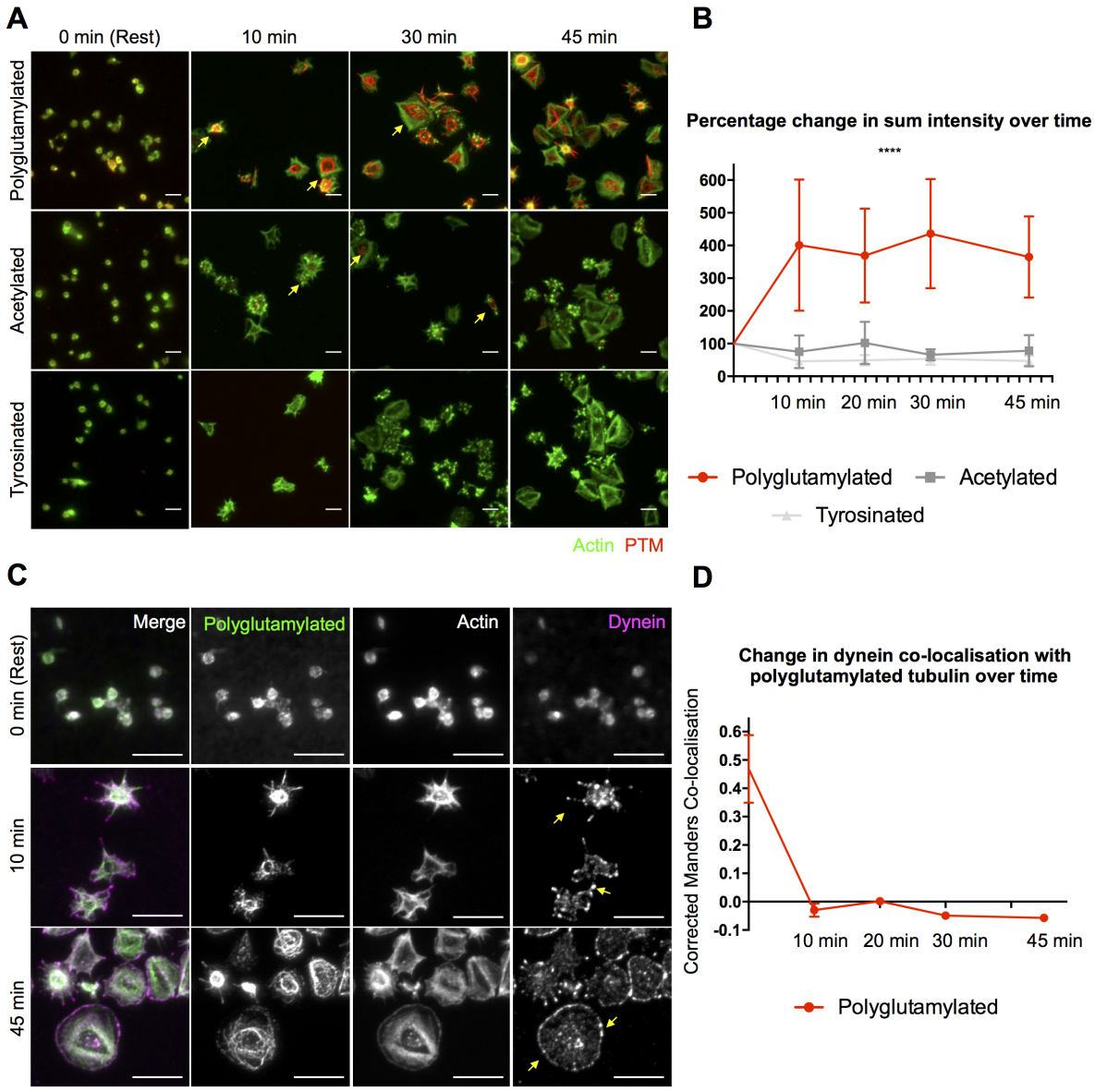

Figure 4.7: **Platelet activation results in hyperglutamylation of the marginal band and a subsequent mobilisation of dynein.** A time course was performed to investigate the rate at which polyglutamylation of the marginal band occurs, and to compare this post-translational modification with acetylation and tyrosination, two modifications previously thought to influence marginal band stability at the resting state (and upon subsequent activation).

Figure 4.7: *Cont.* (A) A comparison of the polyglutamylation, acetylation, and tyrosination of platelets shows a rapid increase in polyglutamylation in as little as 10 minutes of spreading on fibrinogen. Interestingly this increase in polyglutamylation is significantly larger than fold changes in the intensity of acetylated and tyrosinated tubulin (**** p ¡ 0.001), with acetylation of the marginal band only becoming evident at later time courses (20 minutes and on wards as indicated by the yellow arrow in the 30 min image). (B) Quantification of the co-localisation of dynein with polyglutamylated tubulin shows a marked decrease in co-localisation on platelet activation.. One-way ANOVA , S.D. $n = 3$. *10μm scale bar.*

of polymodifications with kinesin-1. Kinesin-1 has been recently reported by Adam *et al.* as a player in platelet secretion[123]. Interestingly a significant reduction in co-localisation with polyglutamylated, monoglycylated, and polyglycylated residues is observed (Figure 4.9). This data suggests an effect on kinesin-1 processivity as a consequence of glutamylation and glycylation.

Unfortunately family A repeatedly declined invitations to provide further samples for analysis. In the absence of patient samples carrying the suspected C-terminal defect, it becomes very difficult to establish whether the loss of these PTMs is driving abnormal platelet production and behaviour. To that end, efforts to reproduce patient mutations in iPSC derived MKs are important, particularly as transfection into cell lines, which will carry a different landscape of modifying enzymes, is unlikely to faithfully recapitulate the highly specific system in platelets and MKs.

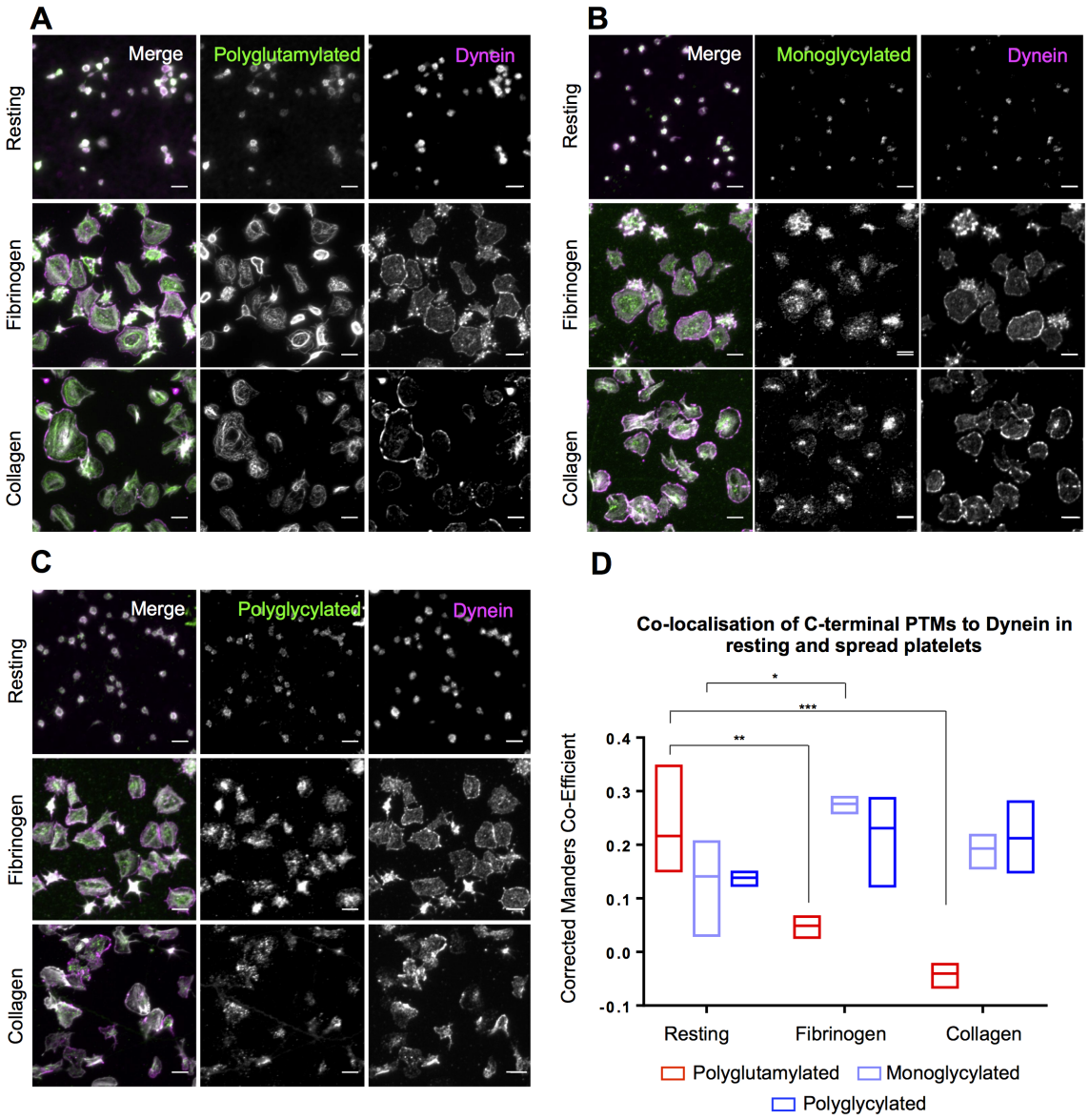

Figure 4.8: **Platelet activation results in a loss of dynein co-localisation with polyglutamylated tubulin, but not with gylycylated residues.** (A) Extensive polyglutamylation of the marginal band is observed in spread cells. In fully spread cells, dynein can be seen to collect on the periphery of the cell, implying that this - end driven motor protein facilitates platelet expansion by mobilising outwards. (B,C) Mono- and polyglycylated residues change distribution on platelet spreading. (D) Quantification of the co-localisation of dynein with each of the three PTMs studied reveals dramatically different relationships (** p = 0.0096). Dynein co-localisation with polyglutamylated residues decreases dramatically on platelet spreading in both fibrinogen and collagen (** p = 0.0082 and *** 0.0004 respectively). Conversely there is a significant increase in the co-localisation of dynein with monoglycylated residues in fibrinogen spread (* p = 0.0251) but not in collagen spread platelets. No significant change in co-localisation with polyglycylated residues is observed. Two-way ANOVA with uncorrected multiple comparisons performed, line of each plot represents mean and min/max values. $n = 3$. *10μm scale bar.*

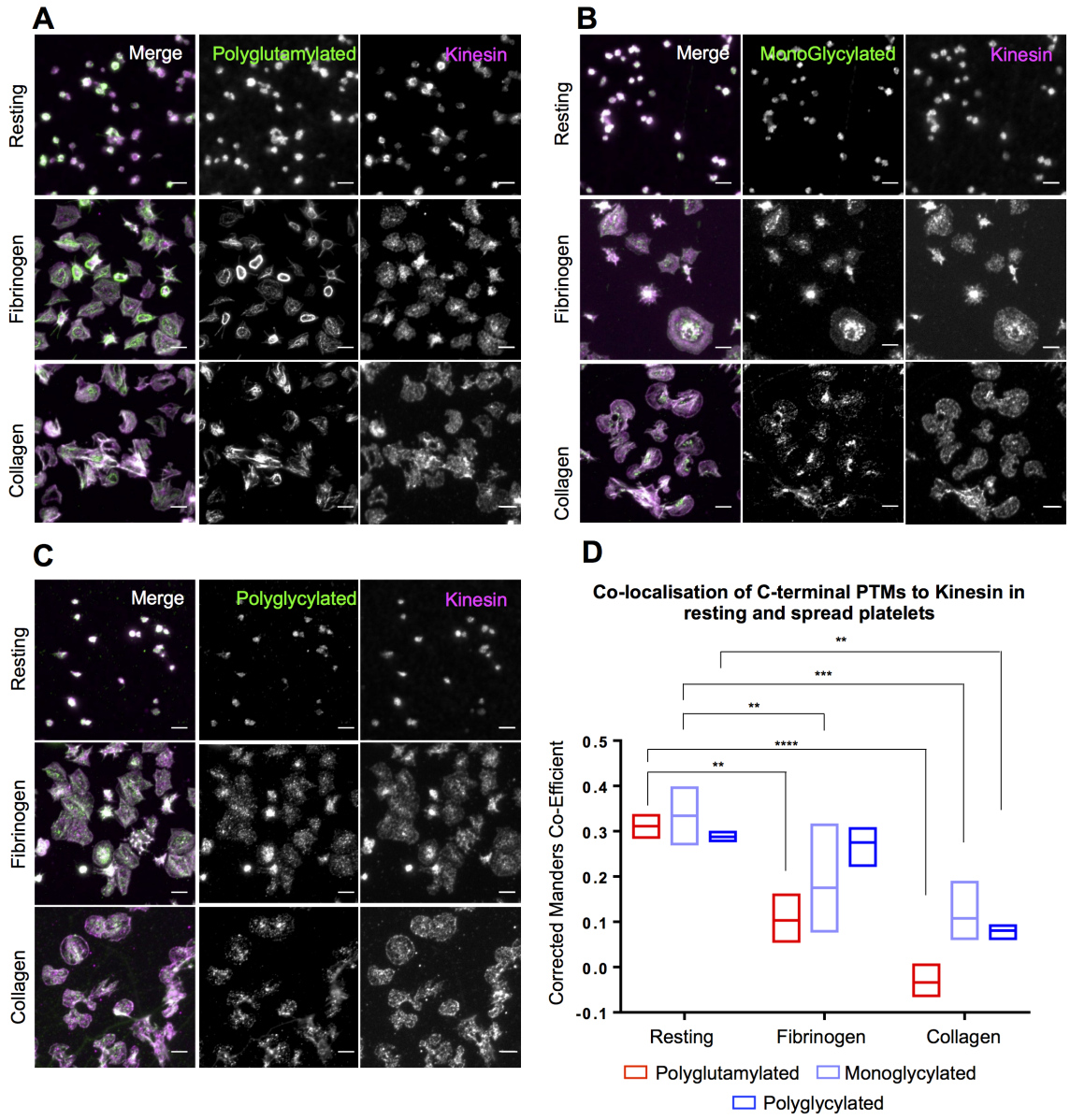

Figure 4.9: **Platelet spreading results in a significant decrease in kinesin localisation with post-translationally modified residues.** (A, B, C) On platelet spreading on fibrinogen and collagen kinesin can be found distributed across the cell, with no substantial co-localisation with polyglutamylated, monoglycylated, or polyglycylated tubulin. (D) The quantification of this co-localisation supports this, showing a significant decrease in co-localisation on both fibrinogen and collagen between kinesin and polyglutamylated (** p = 0.0017, **** p ¡ 0.0001), monoglycylated (** p = 0.0094, *** p = 0.0009) residues. A significant decrease with polyglycylated residues was only observed in fully spread collagen cells (** p = 0.0017). Two-way ANOVA with uncorrected multiple comparisons performed, line of each plot represents mean and min/max values. $n = 3$. *10μm scale bar.*

4.3.3 Designing *TubB1* targeting CRISPRs

One of the goals of this studentship has been adapting genome editing to the targeting of iPSC, and the subsequent generation of MKs from these edited populations. The difficulty in recruiting patients after their initial appointment highlights the importance of being able to produce patient mutations *in vitro* for further mechanistic analysis and study.

To this end, two lines of CRISPR editing were pursued for the study of polymodifications in *TubB1*. The first would be the generation of an mEGFP or mEos 4b CO tagged *TubB1* line. This would offer a means by which to visualise polymodifications in the context of *TubB1* expression, would be a potent tool in studying how specific polyglutamylation and polyglycylation are to the expression of this specific isoform.

A strategy similar to the one pursued for the labelling of *TubA1B* was pursued. As shown in figure 4.10, a donor sequence carrying the insert of choice (mEGFP or mEos 4b CO in this instance) with a linker and flanking homology arms was designed to the N-terminus of *TubB1*. The protein coding region of the first exon of these gene is 57 base pairs long, and is flanked by a 5' UTR and an intron in the 3' direction (Figure 4.10A).

The homology arms and insert sequences were ordered as gBlocks with overhanging arms for gibson assembly, and cloned into an empty, linearised pGem-T-Easy vector as in chapter 2. Upon transformation, clonal isolation, and DNA preparation, clones were verified through an EcoRI digest of the relevant plasmid (Figure 4.10B,C). Control clones were linearised and ran at 3000 base pairs, while successfully assembled plasmids carried an insert which presented as a second band running at approximately 1500-2000 base pairs.

With donor plasmids in hand, guide sequences targeting exon 1 of the *TubB1* gene were generated (Figure 4.11A). Each of the 4 guides was transfected using the Alt-R®method, and as the gene is not expressed, editing efficiency was evaluated through PCR and a T7EI assay.

The T7EI assay generates heteroduplexes of wild type and cleaved DNA which are

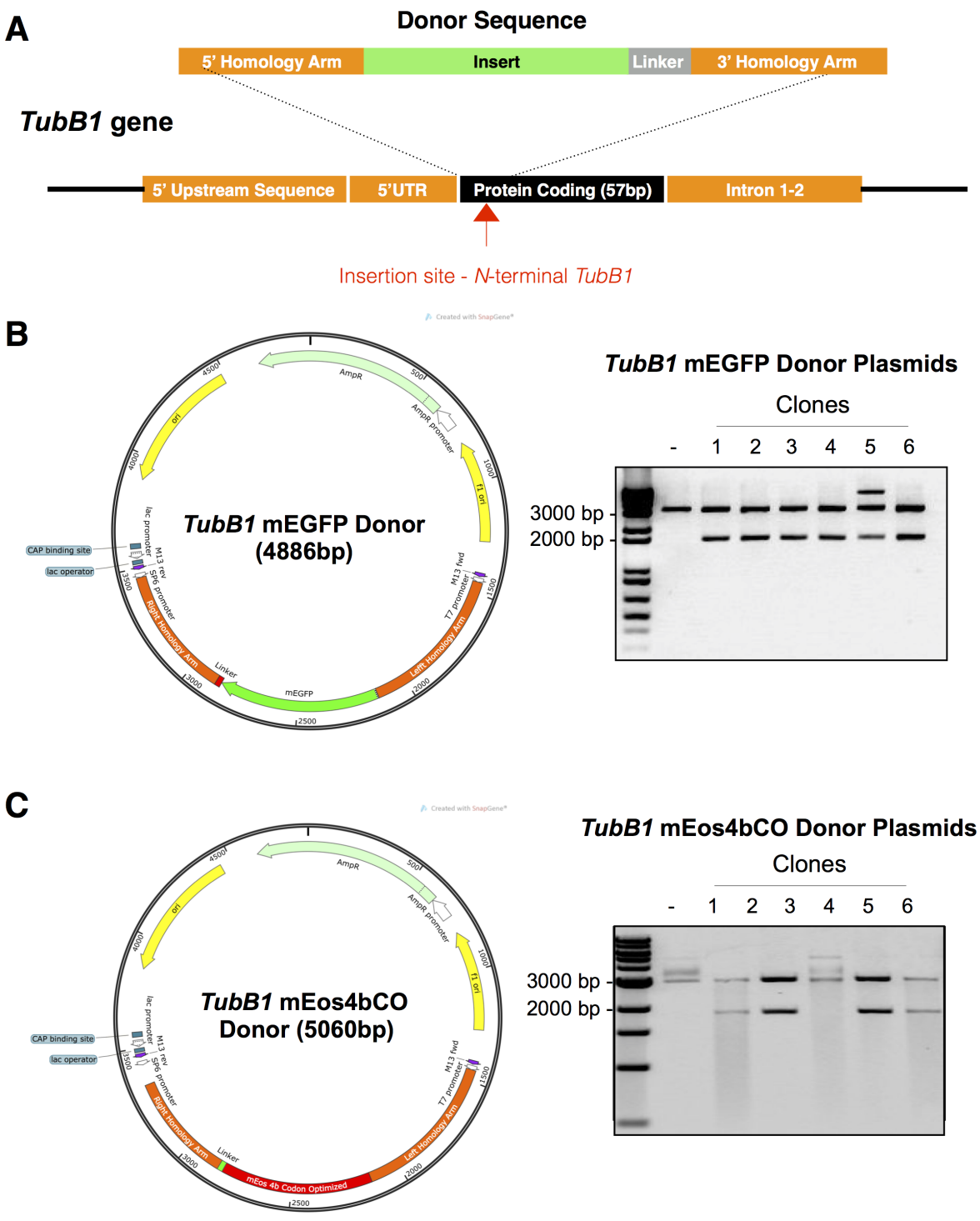

Figure 4.10: **Design and assembly of *TubB1* HDR donors.** A CRISPR mediated HDR knock-in strategy was adapted from previous work to tag the N-terminus of the *TubB1* protein. (A) Homology arms were designed to flank each tag and targeted to the first exon of *TubB1* and sit prior to the 57bp coding region in this exon.

Figure 4.10: *Cont.* (B) An mEGFP carrying donor was generated through gibson assembly into a pGem-T-Easy backbone, colonies were picked and verified with an EcoRI digest before sequencing to confirm a correct assembly. (C) A similar strategy was employed using an mEos 4CO gBlock to generate a *TubB1* HDR donor carrying this photoswitchable reporter. Clones were validated through restriction digest before sequencing.

recognised and cleaved by the T7-Endonuclease. In the presence of edited DNA which is successfully cleaved, a pair of cleavage products can be observed beneath the PCR product in question. As shown in figure 4.11B, a positive control shows two distinctive bands on cleavage. In contrast, untreated iPSC and those transfected with either donor plasmid only or purified Cas9 protein (no guides) show no evidence of cleavage, and so no evidence of heteroduplex formation and thus gene editing.

Cells transfected with each of the guides show some evidence of cleavage, however guide A shows the most substantial level of cleavage, making it the most efficient of the guides tested. As we co-transfect with a donor template, there is evidence of a second band at the predicted molecular weight of an mEGFP-tagged *TubB1*. While there is a faint band at this weight in the donor only control, the increased intensity in the guide treated lane suggests there is some integration of the tag in this population.

Unfortunately on limiting dilution and clonal expansion, none of the 18 screened clones showed any evidence of a tagged *TubB1* gene (Figure 4.11C). This suggests that the amplification of the tagged band in figure 4.11B is an artefact and a consequence of the amplification of the donor template.

Work to generate the tag is ongoing, with experiments testing small molecule inhibitors as discussed in chapter 2 offering the potential for increased integration. Currently a CRISPY treated population is being expanded for validation, as demonstrated in figure 4.11D, this includes testing dsDNA donor templates and the nickase/CRISPY format of this experiment. Primers have been designed to span the region outside of the plasmid sequence to better determine the presence of an insert in these samples.

CRISPR experiments aiming to introduce the two patient SNPs discussed in this work

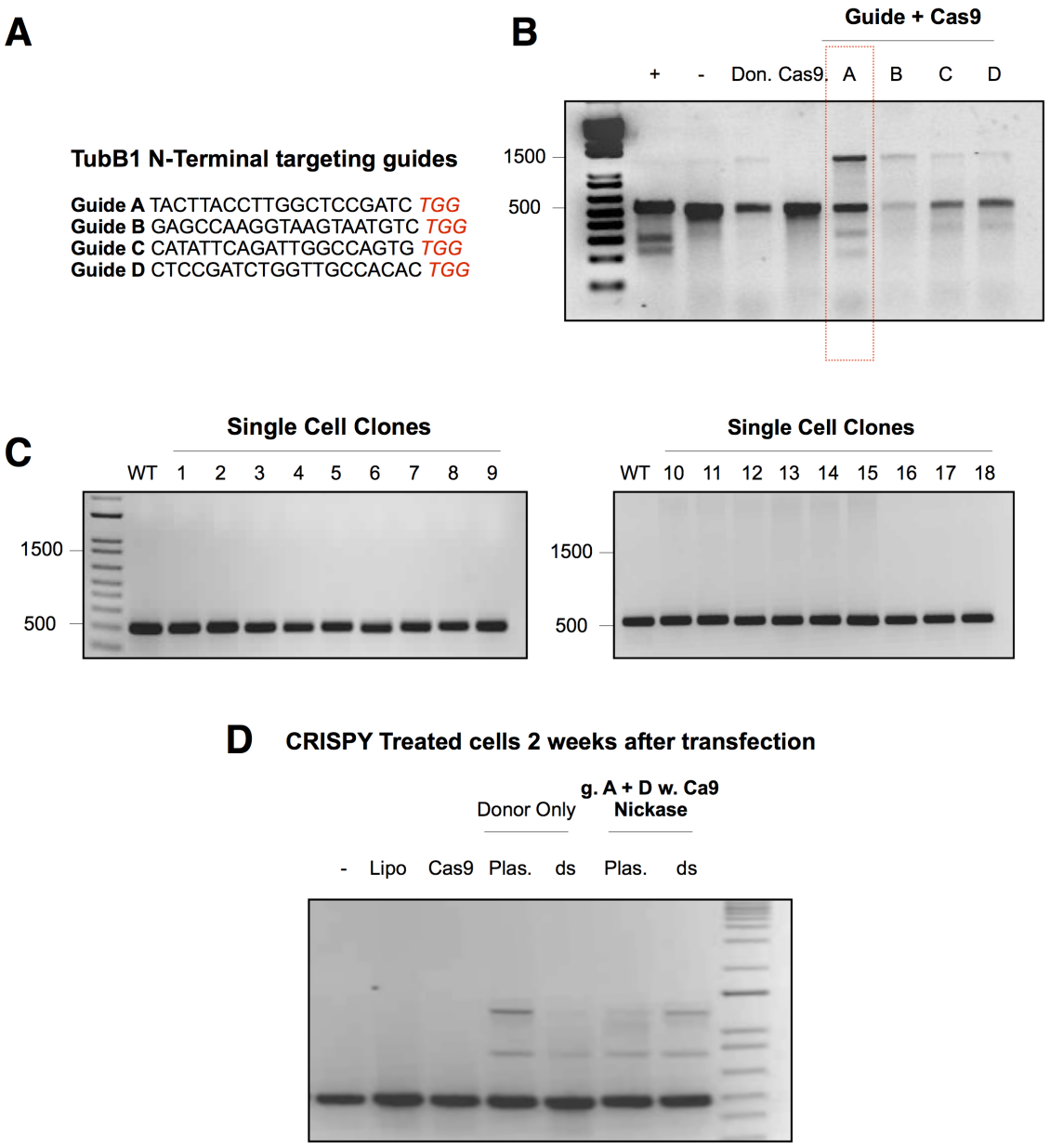

Figure 4.11: **CRISPR mediated labelling of *TubB1* with mEGFP.** Guides targeting exon 1 of *TubB1* were generated and transfected using the Alt-R®system described in chapter 2 into iPSC. (A) Four guides were generated, two of which targeted the forward strand (gA, gB) and reverse strand (gC,gD). (B) Transfected cells were detached 48 hours after transfection and pelleted for DNA extraction. This DNA was then amplified using primers flanking the insert site for analysis through a T7EI assay.

Figure 4.11: *Cont.* The positive control (mismatched DNA duplexes from a Gibco kit) undergo cleavage upon T7EI treatment, resulting in two cleaved bands beneath the 500bp amplicon. Negative, lipofectamine, and Cas9 only treated controls do not appear to have two bands beneath the designed amplicons. Conversely cells treated with guides and Cas9s both demonstrate cleavage bands and an additional heavier band at approximately 1300bp, consistent with a successful integration of the donor mEGFP sequence. Of the guides tested, guide A appeared to cause the most significant cleavage and integration, suggesting this guide is best suited for targeting the *TubB1* locus. (C) Guide A targeted cells were then subject to a limiting dilution before individual clones were expanded and analysed for an mEGFP insertion at this locus. Unfortunately none of the 18 clones carried an mEGFP tagged *TubB1* gene. (D) Current efforts are focussing on increasing the efficiency of insertion to better isolate mEGFP tagged cells.

are ongoing. Two oligonucleotide donors were designed to the target region, one carrying the C>T mutation and the other harbouring the G insert (Figure 4.12). Each of these has been transfected into iPSC with either a wild type Cas9 with a validated guide, or a Cas9 D10A nickase with two guides flanking the target site. T7EI analysis of transfected cells reveal cleavage in each of the samples. Unfortunately the assay is not sensitive enough to detect SNPs, and so clonal isolation will be necessary to determine whether there has been a successful insertion of the donor sequence. Interestingly however, more cleavage is evident in the Gins donor treated sample, which would be consistent with the introduction of a frameshift.

4.3.4 iPSC MKs and polymodifications

Developing iPSC MKs carrying patient mutations is a critical part of determining the effects of our polymodifications of interest, however as these cells are currently being generated, some preliminary work has been performed to investigate whether proplatelet iPSC MKs are positive for the polymodifications observed thus far in wild type platelets.

In iPSC differentiated MKs forming proplatelets, polyglutamylation, monoglycylation, and polyglycylation are observed as shown in figure 4.13. The presence of mono- and polyglycylated residues to this extent is particular interesting, as staining with these antibodies in platelets

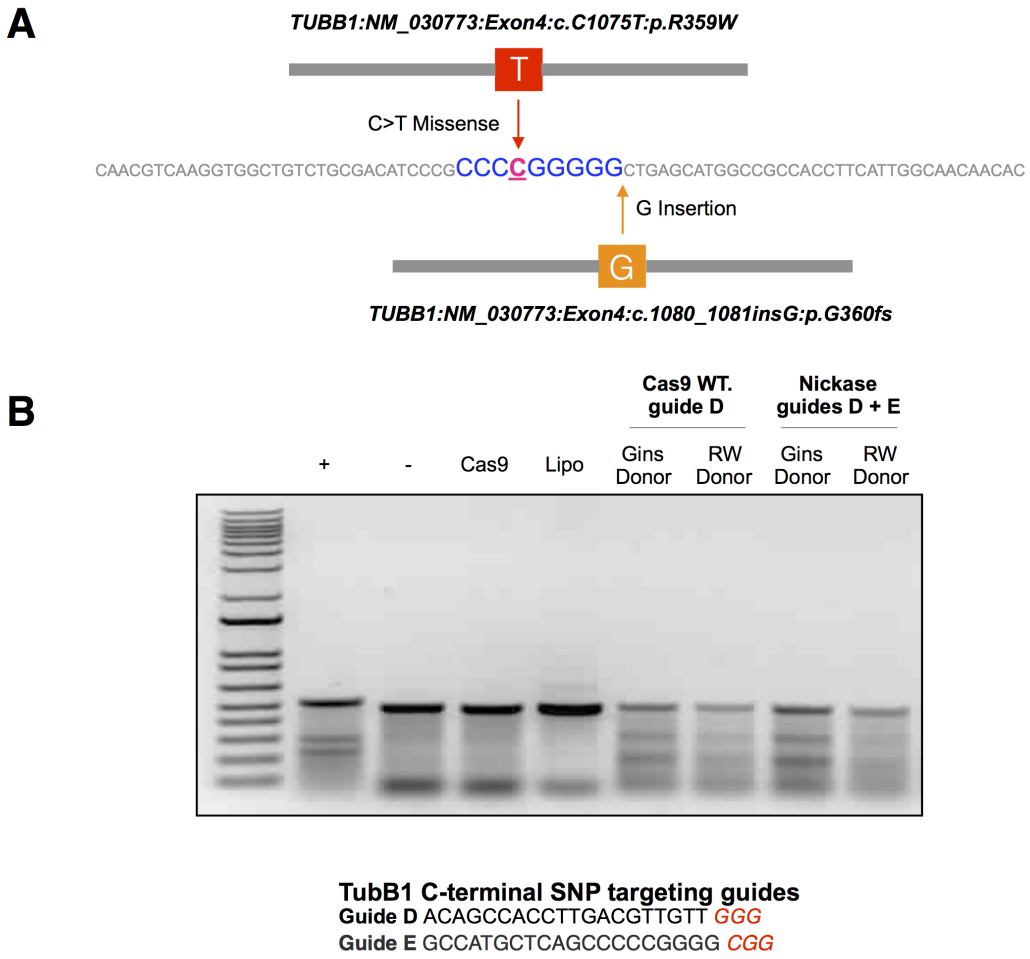

Figure 4.12: **Targeting *TubB1* to introduce patient SNPs in iPSC.** Introducing candidate patient SNPs will be critical to this project and future attempts at disease modelling. (A) Anti-sense oligonucleotides carrying either the C¿T mutation or a G insertion were designed with 45 base pair homology arms. (B) Cells were transfected with guides targeting the SNP region and subject to a T7EI assay after DNA extraction. Untreated, Cas9, and Lipofectamine treated cells do not demonstrate cleavage products after T7EI treatment. Conversely cells treated with either wild type DNA and a previously verified guide D and a Cas9 nickase with two separate guides both demonstrate efficient cleavage after treatment. Transfected cells also carried an RW or G insert donor to facilitate HDR and the introduction of patient SNPs at these loci.

produces markedly different signals and distributions. Notably, CD42 negative cells in these images do not show any evidence of these polymodification, suggesting that they are specific to mature MKs which are likely to express *TubB1*.

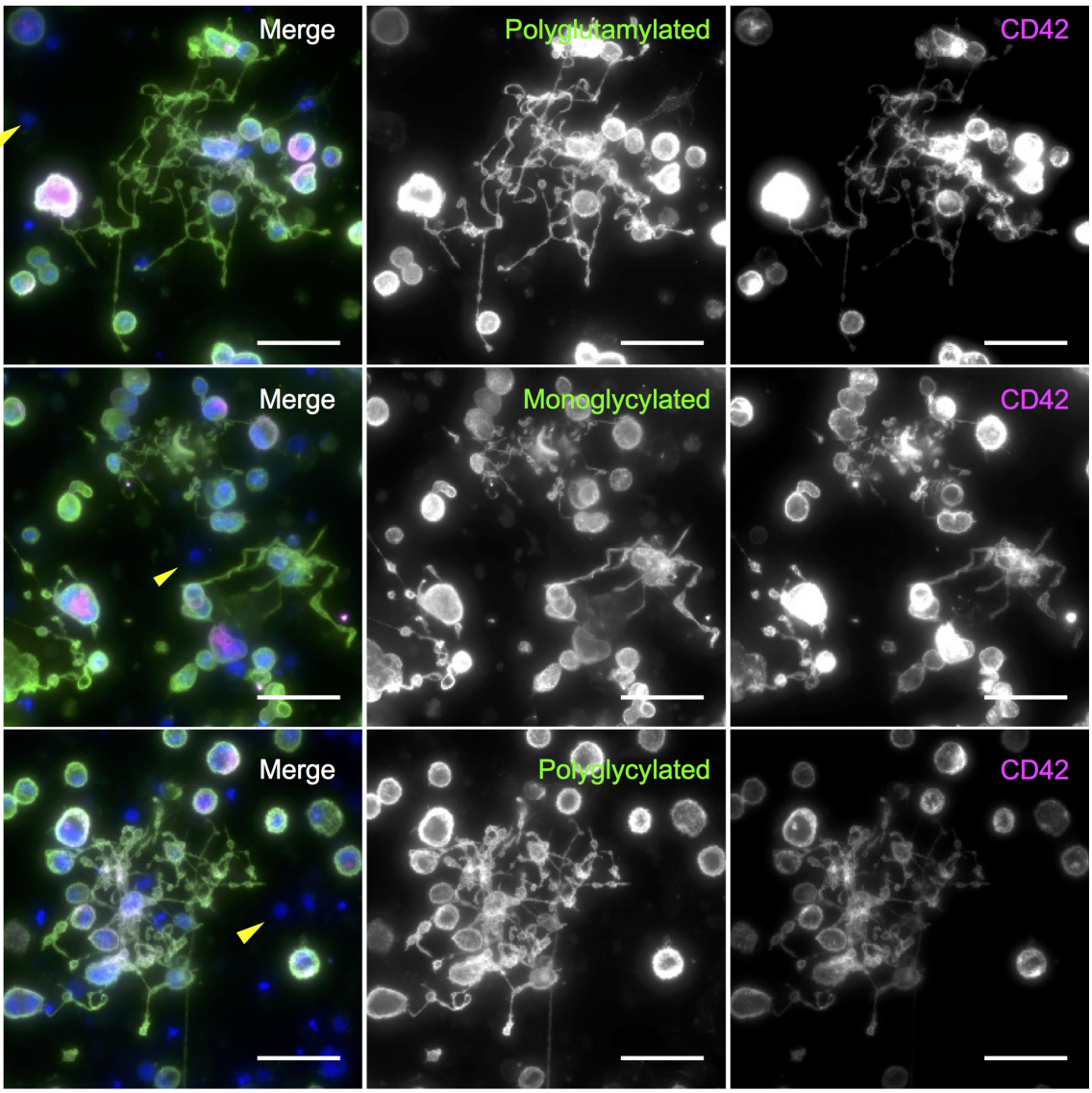

Figure 4.13: **Proplatelet forming human iPSC megakaryocytes demonstrate significant polymodifications specific to MKs and proplatelet forming cells.** Proplatelet forming iPSC MKs show extensive polyglutamylation, mono-, and polyglycylation which appears specific to CD42 positive cells. As indicated by yellow arrows, CD42 negative cells (which are detectable by DAPI staining) are negative for these polymodifications *50μm scale bar.*

4.4 Discussion and Conclusions

The work reported in this chapter represents efforts to translate methodologies developed over the course of this studentship to investigate the mechanistic causes of thrombocytopenia in a particular patient cohort from the GAPP project. At the heart of this is an attempt to improve our understanding of the role of *TubB1* in platelet production and function in the context of the tubulin code.

While the vital role of *TubB1* in platelet production has been shown in both mouse models and human disease case studies, the capacity of this particular isoform to achieve the unique function of both MKs and platelets remains unclear. While the acetylation and tyrosination of the marginal band has been reported in platelets, neither modification has been tied to proplatelet formation or platelet function. Perhaps more importantly, both modifications are ubiquitous and no clear link to *TubB1* has been established.

Of the handful of patient mutations published, the most interesting has been Fiore *et al.*'s report of a C-terminal truncation which exhibits all the hallmarks of a *TubB1* related macrothrombocytopenia[310]. In the context of the tubulin code, this is a very interesting finding as it suggests that post-translational modifications of the β-1 protein are likely to drive its function.

As the C-terminal tail of *TubB1* is rich in glutamate residues, we reasoned that polymodification of these residues is a likely driver of MK and platelet function[113, 264, 273]. This would be consistent with extensive studies in neuronal and ciliated cells which demonstrate critical roles for both glutamylation and glycylation in healthy cell development and function. While not directly comparable to axons or axonemes, there is none the less a parallel between MKs and neurons as demonstrated by Bender *et al.* who show that the dynein driven sliding of MTs first observed in nervous cells is a driver of proplatelet formation.

Our data thus far firstly reports the presence of polyglutamylated and mono/polyglycylated residues in human platelets and megakaryocytes. To our knowledge, this is the first report of

these modifications in this cell lineage.

Secondly we show a critical role for polyglutamylation on platelet activation through spreading on fibrinogen and collagen. Our data shows a significant increase in polyglutamylation on platelet activation which is present across a range of spreading times, and clearly localises to the marginal band on spreading platelets. In resting platelets, there is no evidence of polyglutamylation of this MT ring. We subsequently report a rapid decrease in dynein and kinesin-1 localisation to these polyglutamylated residues, implying that hyperglutamylation of the marginal band is responsible for the mobilisation of motor proteins required for shape change on platelet activation as reported by Diagoura *et al.*.

Evidence of significant polyglutamylation in proplatelet producing MKs suggests that this modification is a likely driver of the dynein driven sliding of MTs required for proplatelet formation. While this data is preliminary, further investigation and the use of iPSC-MKs carrying patient specific mutations is likely to provide invaluable insights into the specificity of these polymodifications to *TubB1*, and the effect of patient impairments on these PTMs.

The role of polyglutamylation in mobilising motor proteins has been previously reported in neuronal cells, and recent work in retinal cells highlights the importance of a ratio between glutamylated and glycylated residues. In excellent work by Sirajuddin *et al.*, polyglutamylation is shown to increase kinesin-1 motility on neuronal β-tubulin in an *in vitro* assay of motor processivity on specific β-tubulin isoforms[313]. In our data, the effects of polyglutamylation include a decrease in the co-localisation of both dynein and kinesin, suggesting that hyperglutamylation increases processivity on platelet activation.

Interestingly our data shows the presence and mobilisation of axonemal dynein in platelets in response to polyglutamylation. To our knowledge, no previous reports have detailed the presence or function of axonemal dynein in non-ciliated cells. That being said, considering that platelets are produced by long processes analogous to cilia in some ways, it is not surprising that axonemal dynein is present and is affected by polyglutamylation of MTs

on activation. The distribution of this dynein subtype on the edge of the cell on platelet spreading strongly suggests that the movement of this motor towards the periphery cell is a driver of cell spreading in this instance.

Interestingly proplatelet forming iPSC MKs also demonstrate extensive glycylation in contrast with platelets stained with these antibodies. In proplatelet forming and mature MKs, these polymodifications are likely to play a markedly different role when compared to platelets. Polyglycylation has not been previously reported in platelets or MKs, in fact work by Rogowski *et al.* suggests an evolutionary redundancy of polyglycylation in humans[285]. While Ikegami *et al.* show that the polyglycylase TTLL10 is functional in its elongase capacity if co-expressed with an appropriate initiase (TTLL8)[314]. This highlights the importance of using a human model for the study of PTMs as the palette of modifying enzymes is species and cell type specific.

Our data shows an increase in polyglycylation, though it is unclear what role this may play in platelet spreading. Surprisingly kinesin-1 poorly colocalises with polyglycylated residues, while dynein shows a significant increase in dynein co-localisation, suggesting that these PTMs may be involved in regulating the motility of this particular motor.

One possible explanation is that glycylated residues act to limit the mobility of dynein, and that while in iPSC Mks this is needed to regulate the extension, twisting, budding and eventual severing of platelets, it is likely that this 'stopping' action is not needed in platelets, where hyperglutamylation of the marginal band is needed to rapidly induce instability and hypermobilisation of dynein.

While the development of fluorescently labelled *TubB1* iPSC lines, and cells carrying patient SNPs is still under way, these measures should provide a substantial insight into the role of these PTMs in the disease state. In the absence of effective inhibitors to TTLLs, these disease models are critical to establishing and completing this new chapter to the tubulin code.

Chapter 5

General Discussion

In this chapter the results presented in this thesis are discussed in the broader context of the literature, with an emphasis on future directions. The key goal of this project was the generation of protocols for CRISPR mediated HDR, and the development of protocols which allow for the differentiation of iPSCs into mature, proplatelet forming MKs. Ultimately the aim is to apply these methods to interrogate the role of novel mutations in causing inherited thrombocytopenias in the GAPP patient cohort.

5.1 Development of CRISPR-PALM

The second chapter of this thesis presented the development of a novel means of single molecule imaging through the knock-in of photoswitchable tags. This method, dubbed CRISPR-PALM, has since been published in Scientific Reports [315].

Interestingly we found a sequence and cell line dependent effect on expression of tagged alleles. While mEGFP was consistently well expressed, mEos 3.2 insertions were substantially down regulated in all three cell lines studies. To our knowledge, no such effect on endogenous expression has been reported. We were also unable to generate homozygous knock-ins at the *TubA1B* locus in any of the cell lines targeted. These findings together suggested, to our

surprise, that endogenous expression of tagged genes is down regulated where the insertion of a tag affects protein function. mEGFP, which is the most monomeric and codon optimised of the many existing fluorescent proteins, is well tolerated, while mEos 3.2, a substantially less monomeric protein is consistently downregulated.

To test the hypothesis that fundamental properties of the fluorophore drive this downregulation, we codon optimised mEos 3.2 and tested a more monomeric variant known as mEos 4b (which we also codon optimised). A comparison of these fluorophores both in their codon optimised and original sequence state reveals that mEos 4b CO is far better expressed than its counterparts, while still not as highly expressed as mEGFP. The improved expression of mEos 4b CO naturally results in more detections, and there improves the quality of single molecule imaging performed through the CRISPR PALM system.

The finding that insertions can so significantly affect allele expression is an extremely important one, with serious implications for knock-in work currently performed. At present little functional validation is performed on knock-in cells, in most published work attention is paid to correctly identifying the genomic change produced by CRISPR experiments. We believe that all insertions should be carefully assessed for downregulation at the protein level, as this is likely to have significant functional consequences. With a wide range of fluorescent proteins available, characterising which tags are best suited to endogenous expression will be critical to achieving the potential of imaging experiments utilising knock-in cells.

From a single molecule perspective a significant finding is that where expression of endogenous mEos 3.2 is high, sufficient labelling is achieved to demonstrate high quality SMLM which is comparable by a number of metrics to dSTORM. Achieving such high level spatial resolution through endogenous expression is a significant finding, particularly for a complex structure like tubulin which demands stringent labelling to achieve high quality SMLM. This data suggests that where the down-regulation of an inserted fluorophore can be limited, endogenous expression is a viable method of single molecule imaging.

Further work must be performed to establish whether these effects are consistent across multiple genes. Currently cell lines carrying these inserts in the Vimentin and CXCR4 genes are being developed and assessed for their expression level and performance in single molecule imaging. Vimentin has been previously reported as a heterozygous knock-in of rsEGFP in U2OS cells by Ratz *et al.*, the authors apply RESOLFT to image these cells and show a marginal improvement in resolution compared to confocal microscopy[206]. Reproducing these results and validating whether mEos 4b CO is better expressed in a separate gene will help confirm our conclusion that tag specific properties are important to the expression of CRISPR edited genes.

Similarly CXCR4 offers an ideal target for testing the potential of CRISPR-PALM in quantifying receptor clustering and distribution on the single molecule level. Single molecule quantification remains a challenging and contentious subject, however the introduction of a system which involves the endogenous expression of a fusion tag may resolve many of the classic issues in this field.

Approaches for both the acquisition of single molecule data and its processing have advanced significantly in the 3 years of this project. High throughput single molecule imaging and automated analysis modalities have substantially advanced the application of these techniques to answering biological questions[316]. The use of lightsheet imaging has further allowed for SMLM in complete volumes with remarkable resolution[196, 180]. The onus is therefore on developing biological samples which can be best used in these modalities. Over-expression results in artifacts and significant cell to cell variation, particularly on the nano-scale. Similarly antibody staining for dSTORM features a number of confounding factors which can influence the quality of results. Developing cell lines with consistent cell to cell expression and with reporters demonstrating defined switching kinetics will be vital to establishing robust, high throughput single molecule imaging.

5.2 Adapting and developing methods of iPSC transfection, CRISPR editing, and differentiation

Developing CRISPR methods for iPSC proved to be significantly more challenging than we had originally thought, particularly in light of our success with HDR in cell lines. While the literature makes it clear that integration efficiencies are substantially lower in stem cells than cell lines, our inability to yield higher rates of insertion, even through the use of small molecules is concerning.

Fortunately independent work by Roberts *et al.* published in late 2017 also reported extremely low targeting of this gene (0.1%) in iPSC [219]. There is a likelihood that targeting this specific allele in iPSC is problematic, and so current work targeting *TubB1* in iPSC is likely to shed further light on the effect of small molecule inhibition on HDR efficiency.

Similarly there remain a number of conditions which have been reported to improve tag insertion in iPSC which we have yet to trial. These include using single stranded donor DNA templates, and altering the design of paired nickase guides to optimally introduce staggered cuts. Unfortunately there is a never ending stream of research on improved HDR in difficult cell types, discerning which approach is likely to suit our needs is an unfortunately tedious approach of trial and error.

The development of a protocol for iPSC differentiation was, fortunately, significantly more successful. Despite early difficulties in determining which media and basement membrane would be best suited to directed differentiation, the resulting protocol efficiently generates populations of mature iPSC MK positive for archetypal markers. More importantly, these cells produce proplatelets *in vitro*, rendering them useful for phenotypically screening knock-ins.

The study of MK and platelet specific genes has suffered due to the absence of human cell lines capable of reproducing this a proplatelet phenotype *in vitro*. As a result, there has been a significant emphasis on mouse models and primary murine megakaryocytes in the

study of platelet and MK defects. The generation of a mouse model is significantly more time consuming and costly than purely *in vitro* work, and has thus limited the throughput and pace of study in this field.

Similarly functional differences between mice and humans are critical when interpreting mouse data in the context of human platelet biology. This is particularly significant when studying the tubulin code, where a cell and species specific landscape of modifying enzymes drives PTMs and their effects.

The current focus of our work is refining the terminal stage of the differentiation protocol in a robust proplatelet assay. The introduction of patient SNPs is a priority, as in the absence of patient material the effects of our candidate mutations can be effectively studied using these differentiation protocols.

5.3 Dissecting the role of novel *TubB1* PTMs in the function of platelets and MKs in the context of inherited thrombocytopenia

The data discussed thus far in chapter 4 reveals a complex co-ordination of motor protein function through post-translational modifications. These findings describe a role for a series of novel PTMs which, to our knowledge, have not yet been reported in platelets and MKs. Polyglutamylation in particular appeaks to play a significant role in destabilising the marginal band by hypermobilising dynein and kinesin on activation, supporting broader work implicating these motors in driving shape change in platelet activation.

The role of glycylation is less obvious, but equally interesting. Monoglycylation is consistent regardless of the activation state of our platelets, however we observe a significant increase in polyglycylation on activation and spreading, suggesting a role for these cells in this process. Interestingly there is an increase in dynein co-localisation with polyglycylated

residues, suggesting that these modifications could alter processivity to regulate motor protein function on spreading. This would be particularly interesting in the context of thrombi, where uncontrolled platelet spreading could be potentially detrimental.

Preliminary evidence shows the presence of polyglutamylation, monoglycylation, and polyglycylation on proplatelet forming cells. In the context of our platelet data and the broader literature, the role of these PTMs on proplatelets is likely to be in the process of microtubule sliding to drive proplatelet extension. However the markedly different distribution of glycylated residues in platelets when compared to their progenitor cells implies a complex role for the regulation of these PTMs in fine tuning the processivity of motor proteins along these extensions.

Unfortunately there is limited data discussing the role of glycylation in microtubule function, particularly in the context of work by Rogowski *et al.* disputing the presence of these polymodifications in humans [285] . A significant flaw in this work is the use of a bone osteosarcoma cell line to determine the effects of candidate tubulin ligases. As Ikegami *et al.* [314] later show, TTLL10 function is specifically as an elongase, and is restored in human cells on co-expression of an initiase (TTLL8). The controversy of these works highlights the need for a species and lineage specific model for the study of polymodifications, as there is likely to be an MK and platelet specific mix of tubulin modifying enzymes working in concert to first extensively polyglutamylate and glycylate proplatelet processes, before removing these modifications in the resting platelet, and finally polyglutamylating the marginal band on platelet activation. As none of the polymodifications rich in iPSC-MKs forming proplatelets are evident in the marginal band of resting platelets, there is likely a critical role for enzymes in reversing these polymodifications in mature, circulating platelets.

As discussed above, the generation of patient specific mutants will be key to linking these observations with the *TubB1* C-terminal tail. Establishing changes in these polymodifications in a diseased stem cell model, and any subsequent functional defects, will be a very exciting

insight into the role of *TubB1* and the tubulin code in megakaryocyte and platelet biology.

Chapter 6

Appendix

Name	Sequence
TubA1B Guide A Top Oligo	CACCGCCTCGACTCTTAGCTTGTCG
TubA1B Guide A Bottom Oligo	AAACCGACAAGCTAAGAGTCGAGGC
TubA1B Guide B Top Oligo	CACCGCCCCGACAAGCTAAGAGTCG
TubA1B Guide B Bottom Oligo	AAACCGACTCTTAGCTTGTCGGGGC
TubA1B Guide C Top Oligo	CACCGACCTCGACTCTTAGCTTGTC
TubA1B Guide C Bottom Oligo	AAACGACAAGCTAAGAGTCGAGGTC
TubA1B Guide D Top Oligo	CACCGATTAGGAGGCGAAGGCGAC
TubA1B Guide D Bottom Oligo	AAACGTCGCCTTCGCCTCCTAATC
TubA1B jPCR Fwd Primer	TAATCTTGCACTCACTCTCTAGGAA
TubA1B jPCR Rev Primer	AGGCATTGTCTGATATTTAGCAAGT
OS280 Primer	CAGGGTTATTGTCTCATGAGCGG

Table 6.1: **Table of *TubA1B* Primer Sequences.**

Name	Sequence
TubB1 SNP Region Primer Fwd	AGACCAGGAACAGCAGCT
TubB1 SNP Region Primer Rev	GGATGGCCGTGTTGTTGC
TubB1 N-Terminal Insertion Fwd	TAATCTTGCACTCACTCTCTAGGAA
TubB1 N-Terminal Insertion Rev	AGGCATTGTCTGATATTTAGCAAGT
TubB1 Sense Donor 1 Gins Fwd	CAACGTCAAGGTCGCTGTCTGCGACATCC CGCCCCGGGGCGCTGAGCATGGCCGCCAC-CTTCATTGGCAACAACACGGCCA

Table 6.2: **Table of *TubB1* Primer Sequences.**

Sequences for homology arms and inserts:

TubA1B Left Homology Arm

tctgcGGCGGCCAGGCCGGGCGCGGAGTGGGCGCGCGGGGCCGGAGGAGGGGCCAG
CGACCGCGGCACCGCCTGTGCCCGCCCGCCCCTCCGCAGCCGCTACTTAAGAGGCTC
CAGCGCCGGCCCCGCCCTAGTGCGTTACTTATCTCGACTCTTAGCTTGTCGGAGACG
GTAACCGGTACCCGGTGTCTGCTCCTGTCGCCTTCGCCTCCTAATCCCTAGCCACT

TubA1B Right Homology Arm

Gtgagtaagccgtgcggctcccggctgctttcagggaagcagggaaaagcgagccggcggggcgctggggccctgtatacagccggg
aagggctggcctcagagccgtccgtttggagggcggaaaacgaggcgagaggccagggcgggagtggtgagacctcggtgtgtgtaaata
gcgggggcccggaaaggtcgaggggcgccaggatttcttctcggactctggaagggatgggggctcgggctgccctccgccgtatccgga
gctctcttttgtcgcgtaactgtgtcctgggtgcggtccctcgagtccccgcagtcctttccagcgcatgcccttactccgccttgggtggacgc
gcgcgcggactcttccagccctcacttcctcttgagcgcgaaaagcgggggtgggaagcagctggagacaaaagcgcgcacgcgcgaccgt
taccttcccgccgctcctgggcgggaaaccgccactgcgctttgcgcatgcgctctgggtcgcgacgcgactagggctacagggcgtgtctcc
tgttaacctgagtgttctttttttgatgaaagcaataagaggactgcggaagagctccctgtcaatgtaccgctctaca

Dronpa

ATGGGAGTGAGTGTGATTAAACCAGACATGAAGATCAAGCTGCGTATGGAAGGC
GCTGTAAATGGACACCCGTTCGCGATTGAAGGAGTTGGCCTTGGGAAGCCTTTCGAG
GGAAAACAGAGTATGGACCTTAAAGTCAAAGAAGGCGGACCTCTGCCTTTCGCCTAT
GACATCTTGACAACTGTGTTCTGTTACGGCAACAGGGTATTCGCCAAATACCCAGAA
AATATAGTAGACTATTTCAAGCAGTCGTTTCCTGAGGGCTACTCTTGGGAACGAAGC
ATGAATTACGAAGACGGGGGCATTTGTAACGCGACAAACGACATAACCCTGGATGGT
GACTGTTATATCTATGAAATTCGATTTGATGGTGTGAACTTTCCTGCCAATGGTCCA
GTTATGCAGAAGAGGACTGTGAAATGGGAGCCATCCACTGAGAAATTGTATGTGCGT
GATGGAGTGCTGAAGGGTGATGTTAACATGGCTCTGTCGCTTGAAGGAGGTGGCCAT
TACCGATGTGACTTCAAAACTACTTATAAAGCTAAGAAGGTTGTCCAGTTGCCAGAC
TATCACTTTGTGGACCACCACATTGAGATTAAAAGCCACGACAAAGATTACAGTAAT
GTTAATCTGCATGAGCACGCCGAAGCGCATTCTGAGCTGCCGAGGCAGGCCAAG

mEGFP

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAG
CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGA
TGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGT
GCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTA
CCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGT
CCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT
GAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA
GGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGT
CTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCA
CAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCT
GAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGC
CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG

mEos 3.2

ATGGTGAGTGCGATTAAGCCAGACATGAAGATCAAACTCCGTATGGAAGGCAACG
TAAACGGGCACCACTTTGTGATCGACGGAGATGGTACAGGCAAGCCTTTTGAGGGAA
AACAGAGTATGGATCTTGAAGTCAAAGAGGGCGGACCTCTGCCTTTTGCCTTTGATA
TCCTGACCACTGCATTCCATTACGGCAACAGGGTATTCGCCAAATATCCAGACAACA
TACAAGACTATTTTAAGCAGTCGTTTCCTAAGGGGTATTCGTGGGAACGAAGCTTGA
CTTTCGAAGACGGGGGCATTTGCAACGCCAGAAACGACATAACAATGGAAGGGGACA
CTTTCTATAATAAAGTTCGATTTTATGGTACCAACTTTCCCGCCAATGGTCCAGTTAT
GCAGAAGAAGACGCTGAAATGGGAGCCCTCCACTGAGAAAATGTATGTGCGTGATGG
AGTGCTGACGGGTGATATTGAGATGGCTTTGTTGCTTGAAGGAAATGCCCATTACCG
ATGTGACTTCAGAACTACTTACAAAGCTAAGGAGAAGGGTGTCAAGTTACCAGGCGC
CCACTTTGTGGACCACTGCATTGAGATTTTAAGCCATGACAAAGATTACAACAAGGT
TAAGCTGTATGAGCATGCTGTTGCTCATTCTGGATTGCCTGACAATGCCAGACGA

mEos 3.2 CO

TCAGCAATAAAACCGGACATGAAGATTAAACTTCGCATGGAGGGCAACGTGAACG

GTCATCATTTCGTCATAGACGGAGACGGCACGGGCAAGCCCTTTGAGGGTAAACAGA
GCATGGACCTGGAAGTGAAGGAGGGTGGTCCTCTCCCCTTTGCGTTTGATATTTTGA
CAACGGCATTCCATTATGGCAATCGGGTATTCGCCAAGTATCCGGACAACATACAAG
ACTATTTTAAACAGTCTTTTCCAAAAGGCTATTCCTGGGAAAGATCCCTCACATTTGA
GGACGGAGGGATATGCAATGCACGAAATGACATAACGATGGAGGGCGATACATTTTA
CAACAAGGTGAGGTTCTACGGGACAAACTTCCCCGCAAATGGCCCAGTTATGCAGAA
GAAGACTTTGAAATGGGAACCTTCAACAGAAAAAATGTACGTCAGGGATGGCGTGCT
TACCGGGGATATTGAGATGGCATTGCTGTTGGAGGGAAACGCTCACTATCGGTGCGA
CTTCCGCACCACCTACAAAGCAAAAGAGAAGGGAGTGAAGTTGCCGGGCGCACATTT
CGTGGACCACTGCATTGAGATCCTCTCCCATGACAAAGATTACAATAAAGTGAAGCT
CTATGAACATGCGGTTGCTCATTCTGGTCTTCCTGACAATGCGCGGCGA

mEos 4b CO

ATGGTTAGTGCGATTAAGCCAGACATGCGGATAAAACTCCGCATGGAGGGGAAC
GTAAACGGGCATCATTTTGTTATAGACGGAGATGGGACCGGCAAACCGTATGAGGGT
AAGCAGACTATGGATCTCGAAGTGAAAGAAGGAGGTCCCTTGCCCTTCGCGTTTGAT
ATACTGACAACTGCCTTCCACTATGGAAATCGCGTTTTTGTGAAGTATCCCGACAAC
ATACAGGACTATTTTAAGCAATCTTTTCCGAAGGGTTATAGCTGGGAGCGCTCATTG
ACTTTTGAGGATGGAGGCATCTGCAATGCACGAAACGATATAACGATGGAGGGGGAT
ACTTTCTACAATAAAGTTAGGTTCTACGGAACAAATTTTCCGGCAAATGGACCTGTC
ATGCAGAAAAAAACCCTTAAATGGGAGCCGAGTACAGAAAAAATGTACGTGCGGGAT
GGGGTGCTGACCGGTGATATTGAAATGGCTCTGCTCTTGGAGGGCAATGCCCATTAT
AGATGCGATTTTAGAACCACCTACAAGGCCAAAGAAAAAGGCGTCAAACTGCCAGGC
GCACATTTTGTAGATCACGCCATTGAGATACTTTCACACGACAAAGACTATAATAAG
GTGAAGCTTTATGAACATGCTGTTGCACACAGCGGTCTTCCTGATAACGCAAGACGA

mEos 4b

ATGGTGAGTGCGATTAAGCCAGACATGAGGATCAAACTCCGTATGGAAGGCAAC
GTAAACGGGCACCACTTTGTGATCGACGGAGATGGTACAGGCAAGCCTTATGAGGGA
AAACAGACCATGGATCTTGAAGTCAAAGAGGGCGGACCTCTGCCTTTTGCCTTTGAT

ATCCTGACCACTGCATTCCATTACGGCAACAGGGTATTCGTGAAATATCCAGACAAC
ATACAAGACTATTTTAAGCAGTCGTTTCCTAAGGGGTATTCGTGGGAACGAAGCTTG
ACTTTCGAAGACGGGGGCATTTGCAATGCCAGAAACGACATAACAATGGAAGGGGAC
ACTTTCTATAATAAAGTTCGATTTTATGGTACCAACTTTCCCGCCAATGGTCCAGTTA
TGCAGAAGAAGACGCTGAAATGGGAGCCCTCCACTGAGAAAATGTATGTGCGTGATG
GAGTGCTGACGGGTGATATTGAGATGGCTTTGTTGCTTGAAGGAAATGCCCATTACC
GATGTGACTTCAGAACTACTTACAAAGCTAAGGAGAAGGGTGTCAAGTTACCAGGCG
CCCACTTTGTGGACCACGCCATTGAGATTTTAAGCCATGACAAAGATTACAACAAGG
TTAAGCTGTATGAGCATGCTGTTGCTCATTCTGGATTGCCTGACAATGCCAGACGA

HaloTag™CO

ATGGCCGAAATCGGCACGGGTTTCCCATTCGACCCCCACTATGTGGAGGTACTGG
GAGAGAGAATGCACTACGTGGACGTAGGTCCCAGAGATGGAACGCCTGTACTGTTTC
TCCATGGGAACCCCACGTCTAGCTATGTCTGGCGGAACATCATTCCTCATGTCGCTC
CGACACATCGATGCATAGCCCCTGATTTGATCGGTATGGGCAAGTCCGACAAACCGG
ACCTTGGGTACTTCTTTGATGACCATGTAAGGTTCATGGATGCTTTCATTGAGGCGC
TCGGACTTGAGGAAGTCGTTTTGGTCATCCATGATTGGGGGAGTGCTTTGGGATTCC
ATTGGGCAAAGCGGAATCCTGAAAGAGTAAAGGGTATCGCGTTCATGGAATTCATTC
GCCCAATACCCACCTGGGATGAGTGGCCTGAGTTTGCGCGAGAGACGTTTCAAGCGT
TCCGAACTACTGACGTAGGCAGGAAACTTATCATTGACCAGAACGTTTTTATAGAGG
GGACATTGCCAATGGGCGTAGTGCGGCCCCTGACGGAGGTCGAAATGGATCACTACC
GGGAGCCCTTTCTGAACCCGGTGGATAGAGAACCTCTCTGGCGCTTCCCTAATGAGC
TCCCAATTGCAGGAGAGCCGGCGAATATAGTTGCGCTCGTCGAAGAGTACATGGATT
GGCTCCACCAAAGCCCCGTACCGAAACTTTTGTTCTGGGGGACCCCAGGAGTGCTCA
TTCCACCTGCCGAGGCTGCCAGGCTCGCGAAATCATTGCCGAATTGCAAAGCCGTTG
ACATTGGGCCTGGCCTGAACCTTCTCCAGGAGGATAACCCTGATCTTATAGGAAGCG
AGATAGCCCGATGGCTTAGCACCCTGGAGATTTCCGGC

TubB1 Left Homology Arm

Gcaggacactgcagagtaacttgcatacagggcgagagagggcatgtgggggagggcgggcagacagtgcctgatgcaggccatcc

cgggacctgggacctgcgaggcccagttgtggagcctcctgctgtggcggggcacggggcttctggggaccaaagtctggaaaaggggccg
ggaagatggacagggaaagcccttggaggtggtctggttacctgggtgagctaggagccacagtcatcataccacggtcactagggccagc
atggtcacctagaagcctgcaaacagtgccagcctccaggccctgtgccccgaggtggcctttaatcttgcactcactctctaggaaatgatG
GGGCAGTATTCTGTGTTGAGGGAGGAAAAACACTCCCTTCCAAAAGCATGACAGGCA
GAAAGCAGAGAAGGGCCAGGACTGGCTGAGGGCGGGGAGCTGGGCCTCTGGGGTGG
ACACACCCTTGGTCACATTGTGAGGGTAGCTTGGTTGGCCAGTCCCACCACTGCAGT
GACCACAGTTGTGTTGGGCTCACACCAGTGAACCGAAGCTCTGGATTCTGAGAGTCT
GAGGATTCCGTGAAGATCTCAGACTTGGGCTCAGAGCAAGG

TubB1 Right Homology Arm

ATGCGTGAAATTGTCCATATTCAGATTGGTCAGTGTGGCAATCAGATCGGAGCCA
AGgtaagtaatgtctggttactaatcctagctttaccacagtccatagaggcggacaaccaaaaaaatacctaagttagcagggaagaca
ataagtctagcagatttaatgtacttgctaaatatcagacaatgcctttgggctttaattcttaaaaaaaaatttcacatgggaaaagaagga
aacaatctttaactaaagagctacaacacacattaagtgcttagagtgttaatttaattaatgttaattttatccatagatacatatgtacaat
acatgccgaccattatattacataatacgtttatcatctttttctttgcccactctgttgcccaggctggagtagagtggcagaatcatggctca
ctatagcctcaacctcctgggctcaagctattctcctgcctcagcctcctgagtagctgggattacaggcatgcaccac

Linker

TCCGGACTCAGATCTCGA

Bibliography

1. Born, G. V. and Cross, M. J. The aggregation of blood platelets. *J Physiol* **168**, 178–95, Aug (1963).

2. Brecher, G. and Cronkite, E. P. Morphology and enumeration of human blood platelets. *J Appl Physiol* **3**(6), 365–77, Dec (1950).

3. Andrews, R. K., López, J. A., and Berndt, M. C. Molecular mechanisms of platelet adhesion and activation. *Int J Biochem Cell Biol* **29**(1), 91–105, Jan (1997).

4. Olsson, A. K. and Cedervall, J. The pro-inflammatory role of platelets in cancer. *Platelets* **29**(6), 569–573, Sep (2018).

5. Dovizio, M., Bruno, A., Contursi, A., Grande, R., and Patrignani, P. Platelets and extracellular vesicles in cancer: diagnostic and therapeutic implications. *Cancer Metastasis Rev* **1**, May (2018).

6. Cedervall, J., Hamidi, A., and Olsson, A.-K. Platelets, nets and cancer. *Thromb Res* **164 Suppl 1**, S148–S152, 04 (2018).

7. Schmitt, A., Guichard, J., Massé, J. M., Debili, N., and Cramer, E. M. Of mice and men: comparison of the ultrastructure of megakaryocytes and platelets. *Exp Hematol* **29**(11), 1295–302, Nov (2001).

8. Franco, A. T., Corken, A., and Ware, J. Platelets at the interface of thrombosis, inflammation, and cancer. *Blood* **126**(5), 582–8, Jul (2015).

9. Nurden, A. T., Nurden, P., Sanchez, M., Andia, I., and Anitua, E. Platelets and wound healing. *Front Biosci* **13**, 3532–48, May (2008).

10. Zeigler, Z., Murphy, S., and Gardner, F. H. Microscopic platelet size and morphology in various hematologic disorders. *Blood* **51**(3), 479–86, Mar (1978).

11. Periayah, M. H., Halim, A. S., and Mat Saad, A. Z. Mechanism action of platelets and crucial blood coagulation pathways in hemostasis. *Int J Hematol Oncol Stem Cell Res* **11**(4), 319–327, Oct (2017).

12. Ho-Tin-Noé, B., Boulaftali, Y., and Camerer, E. Platelets and vascular integrity: how platelets prevent bleeding in inflammation. *Blood* **131**(3), 277–288, Jan (2018).

13. Packham, M. A. Role of platelets in thrombosis and hemostasis. *Can J Physiol Pharmacol* **72**(3), 278–84, Mar (1994).

14. Verhamme, P. and Hoylaerts, M. F. The pivotal role of the endothelium in haemostasis and thrombosis. *Acta Clin Belg* **61**(5), 213–9 (2006).

15. Lever, R. A., Hussain, A., Sun, B. B., Sage, S. O., and Harper, A. G. S. Conventional protein kinase c isoforms differentially regulate adp- and thrombin-evoked ca signalling in human platelets. *Cell Calcium* **58**(6), 577–88, Dec (2015).

16. Smolenski, A. Novel roles of camp/cgmp-dependent signaling in platelets. *J Thromb Haemost* **10**(2), 167–76, Feb (2012).

17. Marcus, A. J., Broekman, M. J., Drosopoulos, J. H., Islam, N., Alyonycheva, T. N., Safier, L. B., Hajjar, K. A., Posnett, D. N., Schoenborn, M. A., Schooley, K. A., Gayle, R. B., and Maliszewski, C. R. The endothelial cell ecto-adpase responsible for inhibition of platelet function is cd39. *J Clin Invest* **99**(6), 1351–60, Mar (1997).

18. Marcus, A. J., Broekman, M. J., Drosopoulos, J. H., Pinsky, D. J., Islam, N., and Maliszewsk, C. R. Inhibition of platelet recruitment by endothelial cell cd39/ecto-adpase: significance for occlusive vascular diseases. *Ital Heart J* **2**(11), 824–30, Nov (2001).

19. Reininger, A. J. Function of von willebrand factor in haemostasis and thrombosis. *Haemophilia* **14 Suppl 5**, 11–26, Nov (2008).

20. Kanaji, S., Fahs, S. A., Shi, Q., Haberichter, S. L., and Montgomery, R. R. Contribution of platelet vs. endothelial vwf to platelet adhesion and hemostasis. *J Thromb Haemost* **10**(8), 1646–52, Aug (2012).

21. Nieswandt, B. and Watson, S. P. Platelet-collagen interaction: is gpvi the central receptor? *Blood* **102**(2), 449–61, Jul (2003).

22. Watson, S. P., Auger, J. M., McCarty, O. J. T., and Pearce, A. C. Gpvi and integrin alphaiib beta3 signaling in platelets. *J Thromb Haemost* **3**(8), 1752–62, Aug (2005).

23. Durrant, T. N., van den Bosch, M. T., and Hers, I. Integrin iib3 outside-in signaling. *Blood* **1**, Aug (2017).

24. Golebiewska, E. M. and Poole, A. W. Platelet secretion: From haemostasis to wound healing and beyond. *Blood Rev* **29**(3), 153–62, May (2015).

25. Furie, B. and Furie, B. C. Thrombus formation in vivo. *J Clin Invest* **115**(12), 3355–62, Dec (2005).

26. Vretenbrant, K., Ramström, S., Bjerke, M., and Lindahl, T. L. Platelet activation via par4 is involved in the initiation of thrombin generation and in clot elasticity development. *Thromb Haemost* **97**(3), 417–24, Mar (2007).

27. Poulter, N. S. and Thomas, S. G. Cytoskeletal regulation of platelet formation: Coordination of f-actin and microtubules. *Int J Biochem Cell Biol* **66**, 69–74, Sep (2015).

28. Machlus, K. R. and Italiano, Jr, J. E. The incredible journey: From megakaryocyte development to platelet formation. *J Cell Biol* **201**(6), 785–96, Jun (2013).

29. Lefrançais, E., Ortiz-Muñoz, G., Caudrillier, A., Mallavia, B., Liu, F., Sayah, D. M., Thornton, E. E., Headley, M. B., David, T., Coughlin, S. R., Krummel, M. F., Leavitt, A. D., Passegué, E., and Looney, M. R. The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors. *Nature* **544**(7648), 105–109, 04 (2017).

30. Borges, I., Sena, I., Azevedo, P., Andreotti, J., Almeida, V., Paiva, A., Santos, G., Guerra, D., Prazeres, P., Mesquita, L. L., Silva, L. S. d. B., Leonel, C., Mintz, A., and Birbrair, A. Lung as a niche for hematopoietic progenitors. *Stem Cell Rev* **13**(5), 567–574, 10 (2017).

31. Laurenti, E. and Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**(7689), 418–426, 01 (2018).

32. Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C., Nowak, D., Boch, T., Hofmann, W.-K., Ho, A. D., Huber, W., Trumpp, A., Essers, M. A. G., and Steinmetz, L. M. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol* **19**(4), 271–281, Apr (2017).

33. Macaulay, I. C., Svensson, V., Labalette, C., Ferreira, L., Hamey, F., Voet, T., Teichmann, S. A., and Cvejic, A. Single-cell rna-sequencing reveals a continuous spectrum of differentiation in hematopoietic cells. *Cell Rep* **14**(4), 966–977, Feb (2016).

34. Lichtman, M. A. The ultrastructure of the hemopoietic environment of the marrow: a review. *Exp Hematol* **9**(4), 391–410, Apr (1981).

35. Lichtman, M. A., Chamberlain, J. K., Simon, W., and Santillo, P. A. Parasinusoidal location of megakaryocytes in marrow: a determinant of platelet release. *Am J Hematol* **4**(4), 303–12 (1978).

36. Scurfield, G. and Radley, J. M. Aspects of platelet formation and release. *Am J Hematol* **10**(3), 285–96 (1981).

37. Tavassoli, M. and Aoki, M. Localization of megakaryocytes in the bone marrow. *Blood Cells* **15**(1), 3–14 (1989).

38. Avecilla, S. T., Hattori, K., Heissig, B., Tejada, R., Liao, F., Shido, K., Jin, D. K., Dias, S., Zhang, F., Hartman, T. E., Hackett, N. R., Crystal, R. G., Witte, L., Hicklin, D. J., Bohlen, P., Eaton, D., Lyden, D., de Sauvage, F., and Rafii, S. Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis. *Nat Med* **10**(1), 64–71, Jan (2004).

39. Avraham, H., Cowley, S., Chi, S. Y., Jiang, S., and Groopman, J. E. Characterization of adhesive interactions between human endothelial cells and megakaryocytes. *J Clin Invest* **91**(6), 2378–84, Jun (1993).

40. Avraham, H., Banu, N., Scadden, D. T., Abraham, J., and Groopman, J. E. Modulation of megakaryocytopoiesis by human basic fibroblast growth factor. *Blood* **83**(8), 2126–32, Apr (1994).

41. Dütting, S., Gaits-Iacovoni, F., Stegner, D., Popp, M., Antkowiak, A., van Eeuwijk, J. M. M., Nurden, P., Stritt, S., Heib, T., Aurbach, K., Angay, O., Cherpokova, D., Heinz, N., Baig, A. A., Gorelashvili, M. G., Gerner, F., Heinze, K. G., Ware, J., Krohne, G., Ruggeri, Z. M., Nurden, A. T., Schulze, H., Modlich, U., Pleines, I., Brakebusch, C., and Nieswandt, B. A cdc42/rhoa regulatory circuit downstream of glycoprotein ib guides transendothelial platelet biogenesis. *Nat Commun* **8**, 15838, Jun (2017).

42. Junt, T., Schulze, H., Chen, Z., Massberg, S., Goerge, T., Krueger, A., Wagner, D. D., Graf, T., Italiano, Jr, J. E., Shivdasani, R. A., and von Andrian, U. H. Dynamic visualization of thrombopoiesis within bone marrow. *Science* **317**(5845), 1767–70, Sep (2007).

43. Stegner, D., vanEeuwijk, J. M. M., Angay, O., Gorelashvili, M. G., Semeniak, D., Pinnecker, J., Schmithausen, P., Meyer, I., Friedrich, M., Dütting, S., Brede, C., Beilhack, A., Schulze, H., Nieswandt, B., and Heinze, K. G. Thrombopoiesis is spatially regulated by the bone marrow vasculature. *Nat Commun* **8**(1), 127, 07 (2017).

44. Thon, J. N., Dykstra, B. J., and Beaulieu, L. M. Platelet bioreactor: accelerated evolution of design and manufacture. *Platelets* **28**(5), 472–477, Jul (2017).

45. Stritt, S. and Nieswandt, B. In vitro platelets in sight. *Blood* **124**(12), 1849–50, Sep (2014).

46. Martinez, A. F., McMahon, R. D., Horner, M., and Miller, W. M. A uniform-shear rate microfluidic bioreactor for real-time study of proplatelet formation and rapidly-released platelets. *Biotechnol Prog* **33**(6), 1614–1629, 11 (2017).

47. Focosi, D. and Amabile, G. Induced pluripotent stem cell-derived red blood cells and platelet concentrates: From bench to bedside. *Cells* **7**(1), Dec (2017).

48. Eicke, D., Baigger, A., Schulze, K., Latham, S. L., Halloin, C., Zweigerdt, R., Guzman, C. A., Blasczyk, R., and Figueiredo, C. Large-scale production of megakaryocytes in microcarrier-supported stirred suspension bioreactors. *Sci Rep* **8**(1), 10146, Jul (2018).

49. Blin, A., Le Goff, A., Magniez, A., Poirault-Chassac, S., Teste, B., Sicot, G., Nguyen, K. A., Hamdi, F. S., Reyssat, M., and Baruch, D. Microfluidic model of the platelet-generating organ: beyond bone marrow biomimetics. *Sci Rep* **6**, 21700, Feb (2016).

50. Debili, N., Coulombel, L., Croisille, L., Katz, A., Guichard, J., Breton-Gorius, J., and Vainchenker, W. Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow. *Blood* **88**(4), 1284–96, Aug (1996).

51. Hunt, P. A bipotential megakaryocyte/erythrocyte progenitor cell: the link between erythropiesis and megakaryopoiesis becomes stronger. *J Lab Clin Med* **125**(3), 303–4, Mar (1995).

52. Nakeff, A. and Daniels-McQueen, S. In vitro colony assay for a new class of megakaryocyte precursor: colony-forming unit megakaryocyte (cfu-m). *Proc Soc Exp Biol Med* **151**(3), 587–90, Mar (1976).

53. Robert, A., Cortin, V., Garnier, A., and Pineault, N. Megakaryocyte and platelet production from human cord blood stem cells. *Methods Mol Biol* **788**, 219–47 (2012).

54. Thiele, J., Wagner, S., Dienemann, D., Wienhold, S., Fischer, R., and Stein, H. Megakaryocyte precursors (promegakaryoblasts and megakaryoblasts) in the normal human bone marrow. an immunohistochemical and morphometric study on routinely processed trephine biopsies. *Anal Quant Cytol Histol* **12**(4), 285–9, Aug (1990).

55. Szalai, G., LaRue, A. C., and Watson, D. K. Molecular mechanisms of megakaryopoiesis. *Cell Mol Life Sci* **63**(21), 2460–76, Nov (2006).

56. Psaila, B., Lyden, D., and Roberts, I. Megakaryocytes, malignancy and bone marrow vascular niches. *J Thromb Haemost* **10**(2), 177–88, Feb (2012).

57. Kaushansky, K. Megakaryocytes their precursors and their progeny. *Trends Cardiovasc Med* **6**(8), 261–4, Nov (1996).

58. Kaushansky, K., Broudy, V. C., Lin, N., Jorgensen, M. J., McCarty, J., Fox, N., Zucker-Franklin, D., and Lofton-Day, C. Thrombopoietin, the mp1 ligand, is essential for full megakaryocyte development. *Proc Natl Acad Sci U S A* **92**(8), 3234–8, Apr (1995).

59. Kaushansky, K. Thrombopoietin: the primary regulator of megakaryocyte and platelet production. *Thromb Haemost* **74**(1), 521–5, Jul (1995).

60. Kaushansky, K. Historical review: megakaryopoiesis and thrombopoiesis. *Blood* **111**(3), 981–6, Feb (2008).

61. Hitchcock, I. S. and Kaushansky, K. Thrombopoietin from beginning to end. *Br J Haematol* **165**(2), 259–68, Apr (2014).

62. Ito, T., Ishida, Y., Kashiwagi, R., and Kuriya, S. Recombinant human c-mpl ligand is not a direct stimulator of proplatelet formation in mature human megakaryocytes. *Br J Haematol* **94**(2), 387–90, Aug (1996).

63. Lok, S., Kaushansky, K., Holly, R. D., Kuijper, J. L., Lofton-Day, C. E., Oort, P. J., Grant, F. J., Heipel, M. D., Burkhead, S. K., and Kramer, J. M. Cloning and expression of murine thrombopoietin cdna and stimulation of platelet production in vivo. *Nature* **369**(6481), 565–8, Jun (1994).

64. Sohma, Y., Akahori, H., Seki, N., Hori, T., Ogami, K., Kato, T., Shimada, Y., Kawamura, K., and Miyazaki, H. Molecular cloning and chromosomal localization of the human thrombopoietin gene. *FEBS Lett* **353**(1), 57–61, Oct (1994).

65. Bartley, T. D., Bogenberger, J., Hunt, P., Li, Y. S., Lu, H. S., Martin, F., Chang, M. S., Samal, B., Nichol, J. L., and Swift, S. Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor mpl. *Cell* **77**(7), 1117–24, Jul (1994).

66. Gurney, A. L., Carver-Moore, K., de Sauvage, F. J., and Moore, M. W. Thrombocytopenia in c-mpl-deficient mice. *Science* **265**(5177), 1445–7, Sep (1994).

67. de Sauvage, F. J., Hass, P. E., Spencer, S. D., Malloy, B. E., Gurney, A. L., Spencer, S. A., Darbonne, W. C., Henzel, W. J., Wong, S. C., and Kuang, W. J. Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl ligand. *Nature* **369**(6481), 533–8, Jun (1994).

68. Zeigler, F. C., de Sauvage, F., Widmer, H. R., Keller, G. A., Donahue, C., Schreiber, R. D., Malloy, B., Hass, P., Eaton, D., and Matthews, W. In vitro megakaryocytopoietic and thrombopoietic activity of c-mpl ligand (tpo) on purified murine hematopoietic stem cells. *Blood* **84**(12), 4045–52, Dec (1994).

69. Machlus, K. R., Johnson, K. E., Kulenthirarajan, R., Forward, J. A., Tippy, M. D., Soussou, T. S., El-Husayni, S. H., Wu, S. K., Wang, S., Watnick, R. S., Italiano, Jr, J. E., and Battinelli, E. M. Ccl5 derived from platelets increases megakaryocyte proplatelet formation. *Blood* **127**(7), 921–6, Feb (2016).

70. Moreau, T., Evans, A. L., Vasquez, L., Tijssen, M. R., Yan, Y., Trotter, M. W., Howard, D., Colzani, M., Arumugam, M., Wu, W. H., Dalby, A., Lampela, R., Bouet, G., Hobbs, C. M., Pask, D. C., Payne, H., Ponomaryov, T., Brill, A., Soranzo, N., Ouwehand, W. H., Pedersen, R. A., and Ghevaert, C. Large-scale production of megakaryocytes from human pluripotent stem cells by chemically defined forward programming. *Nat Commun* **7**, 11208, Apr (2016).

71. Tijssen, M. R. and Ghevaert, C. Transcription factors in late megakaryopoiesis and related platelet disorders. *J Thromb Haemost* **11**(4), 593–604, Apr (2013).

72. Huang, Z., Dore, L. C., Li, Z., Orkin, S. H., Feng, G., Lin, S., and Crispino, J. D. Gata-2 reinforces megakaryocyte development in the absence of gata-1. *Mol Cell Biol* **29**(18), 5168–80, Sep (2009).

73. Schlegelberger, B. and Heller, P. G. Runx1 deficiency (familial platelet disorder with predisposition to myeloid leukemia, fpdmm). *Semin Hematol* **54**(2), 75–80, 04 (2017).

74. Schmit, J. M., Turner, D. J., Hromas, R. A., Wingard, J. R., Brown, R. A., Li, Y., Li, M. M., Slayton, W. B., and Cogle, C. R. Two novel runx1 mutations in a patient with congenital thrombocytopenia that evolved into a high grade myelodysplastic syndrome. *Leuk Res Rep* **4**(1), 24–7 (2015).

75. Bluteau, D., Glembotsky, A. C., Raimbault, A., Balayn, N., Gilles, L., Rameau, P., Nurden, P., Alessi, M. C., Debili, N., Vainchenker, W., Heller, P. G., Favier, R., and Raslova, H. Dysmegakaryopoiesis of fpd/aml pedigrees with constitutional runx1 mutations is linked to myosin ii deregulated expression. *Blood* **120**(13), 2708–18, Sep (2012).

76. Favier, R., Jondeau, K., Boutard, P., Grossfeld, P., Reinert, P., Jones, C., Bertoni, F., and Cramer, E. M. Paris-trousseau syndrome : clinical, hematological, molecular data of ten new cases. *Thromb Haemost* **90**(5), 893–7, Nov (2003).

77. Schlaeger, T. M., Mikkola, H. K. A., Gekas, C., Helgadottir, H. B., and Orkin, S. H. Tie2cre-mediated gene ablation defines the stem-cell leukemia gene (scl/tal1)-dependent window during hematopoietic stem-cell development. *Blood* **105**(10), 3871–4, May (2005).

78. Chagraoui, H., Kassouf, M., Banerjee, S., Goardon, N., Clark, K., Atzberger, A., Pearce, A. C., Skoda, R. C., Ferguson, D. J. P., Watson, S. P., Vyas, P., and Porcher, C. Scl-mediated regulation of the cell-cycle regulator p21 is critical for murine megakaryopoiesis. *Blood* **118**(3), 723–35, Jul (2011).

79. Rubinstein, J. D., Elagib, K. E., and Goldfarb, A. N. Cyclic amp signaling inhibits megakaryocytic differentiation by targeting transcription factor 3 (e2a) cyclin-dependent kinase inhibitor 1a (cdkn1a) transcriptional axis. *J Biol Chem* **287**(23), 19207–15, Jun (2012).

80. Medves, S., Noël, L. A., Montano-Almendras, C. P., Albu, R. I., Schoemans, H., Constantinescu, S. N., and Demoulin, J.-B. Multiple oligomerization domains of kank1-pdgfr are required for jak2-independent hematopoietic cell proliferation and signaling via stat5 and erk. *Haematologica* **96**(10), 1406–14, Oct (2011).

81. Shivdasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J., and Orkin, S. H. Transcription factor nf-e2 is required for platelet formation independent of the actions of thrombopoietin/mgdf in megakaryocyte development. *Cell* **81**(5), 695–704, Jun (1995).

82. Lecine, P. and Shivdasani, R. A. Cellular and molecular biology of megakaryocyte differentiation in the absence of lineage-restricted transcription factors. *Stem Cells* **16 Suppl 2**, 91–5 (1998).

83. Schwer, H. D., Lecine, P., Tiwari, S., Italiano, Jr, J. E., Hartwig, J. H., and Shivdasani, R. A. A lineage-restricted and divergent beta-tubulin isoform is essential for the biogenesis, structure and function of blood platelets. *Curr Biol* **11**(8), 579–86, Apr (2001).

84. Lecine, P., Italiano, Jr, J. E., Kim, S. W., Villeval, J. L., and Shivdasani, R. A. Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor nf-e2. *Blood* **96**(4), 1366–73, Aug (2000).

85. Ebbe, S. and Stohlman, Jr, F. Megakaryocytopoiesis in the rat. *Blood* **26**, 20–35, Jul (1965).

86. Ebbe, S. Biology of megakaryocytes. *Prog Hemost Thromb* **3**, 211–29 (1976).

87. Therman, E., Sarto, G. E., and Stubblefield, P. A. Endomitosis: a reappraisal. *Hum Genet* **63**(1), 13–8 (1983).

88. Raslova, H., Roy, L., Vourc'h, C., Le Couedic, J. P., Brison, O., Metivier, D., Feunteun, J., Kroemer, G., Debili, N., and Vainchenker, W. Megakaryocyte polyploidization is associated with a functional gene amplification. *Blood* **101**(2), 541–4, Jan (2003).

89. Ravid, K., Lu, J., Zimmet, J. M., and Jones, M. R. Roads to polyploidy: the megakaryocyte example. *J Cell Physiol* **190**(1), 7–20, Jan (2002).

90. Nagata, Y., Muro, Y., and Todokoro, K. Thrombopoietin-induced polyploidization of bone marrow megakaryocytes is due to a unique regulatory mechanism in late mitosis. *J Cell Biol* **139**(2), 449–57, Oct (1997).

91. Vitrat, N., Cohen-Solal, K., Pique, C., Le Couedic, J. P., Norol, F., Larsen, A. K., Katz, A., Vainchenker, W., and Debili, N. Endomitosis of human megakaryocytes are due to abortive mitosis. *Blood* **91**(10), 3711–23, May (1998).

92. Zhang, Y., Wang, Z., and Ravid, K. The cell cycle in polyploid megakaryocytes is associated with reduced activity of cyclin b1-dependent cdc2 kinase. *J Biol Chem* **271**(8), 4266–72, Feb (1996).

93. Kautz, J. and De Marsh, Q. B. Electron microscopy of sectioned blood and bone marrow elements. *Rev Hematol* **10**(2), 314–23; discussion, 324–44 (1955).

94. Schulze, H., Korpal, M., Hurov, J., Kim, S.-W., Zhang, J., Cantley, L. C., Graf, T., and Shivdasani, R. A. Characterization of the megakaryocyte demarcation membrane system and its role in thrombopoiesis. *Blood* **107**(10), 3868–75, May (2006).

95. Eckly, A., Heijnen, H., Pertuy, F., Geerts, W., Proamer, F., Rinckel, J.-Y., Léon, C., Lanza, F., and Gachet, C. Biogenesis of the demarcation membrane system (dms) in megakaryocytes. *Blood* **123**(6), 921–30, Feb (2014).

96. Daimon, T. and Gotoh, Y. Cytochemical evidence of the origin of the dense tubular system in the mouse platelet. *Histochemistry* **76**(2), 189–96 (1982).

97. Gerrard, J. M., White, J. G., Rao, G. H., and Townsend, D. Localization of platelet prostaglandin production in the platelet dense tubular system. *Am J Pathol* **83**(2), 283–98, May (1976).

98. Whiteheart, S. W. Platelet granules: surprise packages. *Blood* **118**(5), 1190–1, Aug (2011).

99. Handagama, P., Rappolee, D. A., Werb, Z., Levin, J., and Bainton, D. F. Platelet alpha-granule fibrinogen, albumin, and immunoglobulin g are not synthesized by rat and mouse megakaryocytes. *J Clin Invest* **86**(4), 1364–8, Oct (1990).

100. Harrison, P. and Cramer, E. M. Platelet alpha-granules. *Blood Rev* **7**(1), 52–62, Mar (1993).

101. Hughes, M., Hayward, C. P., Horsewood, P., Warkentin, T. E., and Kelton, J. G. Measurement of endogenous and exogenous alpha-granular platelet proteins in patients with immune and nonimmune thrombocytopenia. *Br J Haematol* **106**(3), 762–70, Sep (1999).

102. Italiano, Jr, J. E., Richardson, J. L., Patel-Hett, S., Battinelli, E., Zaslavsky, A., Short, S., Ryeom, S., Folkman, J., and Klement, G. L. Angiogenesis is regulated by a novel mechanism: pro- and antiangiogenic proteins are organized into separate platelet alpha granules and differentially released. *Blood* **111**(3), 1227–33, Feb (2008).

103. Kamykowski, J., Carlton, P., Sehgal, S., and Storrie, B. Quantitative immunofluorescence mapping reveals little functional coclustering of proteins within platelet -granules. *Blood* **118**(5), 1370–3, Aug (2011).

104. Djaldetti, M., Fishman, P., Bessler, H., and Notti, I. Sem observations on the mechanism of platelet release from megakaryocytes. *Thromb Haemost* **42**(2), 611–20, Aug (1979).

105. Ihzumi, T., Hattori, A., Sanada, M., and Muto, M. Megakaryocyte and platelet formation: a scanning electron microscope study in mouse spleen. *Arch Histol Jpn* **40**(4), 305–20, Sep (1977).

106. Shaklai, M. and Tavassoli, M. Demarcation membrane system in rat megakaryocyte and the mechanism of platelet formation: a membrane reorganization process. *J Ultrastruct Res* **62**(3), 270–85, Mar (1978).

107. Kosaki, G. In vivo platelet production from mature megakaryocytes: does platelet release occur via proplatelets? *Int J Hematol* **81**(3), 208–19, Apr (2005).

108. Hartwig, J. and Italiano, Jr, J. The birth of the platelet. *J Thromb Haemost* **1**(7), 1580–6, Jul (2003).

109. Radley, J. M. and Hartshorn, M. A. Megakaryocyte fragments and the microtubule coil. *Blood Cells* **12**(3), 603–14 (1987).

110. Vijey, P., Posorske, B., and Machlus, K. R. In vitro culture of murine megakaryocytes from fetal liver-derived hematopoietic stem cells. *Platelets* **ePub**, 1–6, Jul (2018).

111. Radley, J. M. and Haller, C. J. The demarcation membrane system of the megakaryocyte: a misnomer? *Blood* **60**(1), 213–9, Jul (1982).

112. Italiano, Jr, J. E., Lecine, P., Shivdasani, R. A., and Hartwig, J. H. Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes. *J Cell Biol* **147**(6), 1299–312, Dec (1999).

113. Gadadhar, S., Bodakuntla, S., Natarajan, K., and Janke, C. The tubulin code at a glance. *J Cell Sci* **130**(8), 1347–1353, Apr (2017).

114. Chakraborti, S., Natarajan, K., Curiel, J., Janke, C., and Liu, J. The emerging role of the tubulin code: From the tubulin molecule to neuronal function and disease. *Cytoskeleton (Hoboken)* **73**(10), 521–550, Oct (2016).

115. Yang, H., Ganguly, A., Yin, S., and Cabral, F. Megakaryocyte lineage-specific class vi -tubulin suppresses microtubule dynamics, fragments microtubules, and blocks cell division. *Cytoskeleton (Hoboken)* **68**(3), 175–87, Mar (2011).

116. Wang, D., Villasante, A., Lewis, S. A., and Cowan, N. J. The mammalian beta-tubulin repertoire: hematopoietic expression of a novel, heterologous beta-tubulin isotype. *J Cell Biol* **103**(5), 1903–10, Nov (1986).

117. Gremmel, T., Frelinger, 3rd, A. L., and Michelson, A. D. Platelet physiology. *Semin Thromb Hemost* **42**(3), 191–204, Apr (2016).

118. Bender, M., Thon, J. N., Ehrlicher, A. J., Wu, S., Mazutis, L., Deschmann, E., Sola-Visner, M., Italiano, J. E., and Hartwig, J. H. Microtubule sliding drives proplatelet elongation and is dependent on cytoplasmic dynein. *Blood* **125**(5), 860–8, Jan (2015).

119. Tanenbaum, M. E., Vale, R. D., and McKenney, R. J. Cytoplasmic dynein crosslinks and slides anti-parallel microtubules using its two motor domains. *Elife* **2**, e00943, Sep (2013).

120. Roy, S. Dynein's life in the slow lane. *Neuron* **90**(5), 907–9, 06 (2016).

121. Twelvetrees, A. E., Pernigo, S., Sanger, A., Guedes-Dias, P., Schiavo, G., Steiner, R. A., Dodding, M. P., and Holzbaur, E. L. F. The dynamic localization of cytoplasmic dynein in neurons is driven by kinesin-1. *Neuron* **90**(5), 1000–15, 06 (2016).

122. Richardson, J. L., Shivdasani, R. A., Boers, C., Hartwig, J. H., and Italiano, Jr, J. E. Mechanisms of organelle transport and capture along proplatelets during platelet production. *Blood* **106**(13), 4066–75, Dec (2005).

123. Adam, F., Kauskot, A., Kurowska, M., Goudin, N., Munoz, I., Bordet, J.-C., Huang, J.-D., Bryckaert, M., Fischer, A., Borgel, D., de Saint Basile, G., Christophe, O. D., and Ménasché, G. Kinesin-1 is a new actor involved in platelet secretion and thrombus stability. *Arterioscler Thromb Vasc Biol* **38**(5), 1037–1051, May (2018).

124. Kunishima, S., Kojima, T., Matsushita, T., Tanaka, T., Tsurusawa, M., Furukawa, Y., Nakamura, Y., Okamura, T., Amemiya, N., Nakayama, T., Kamiya, T., and Saito, H. Mutations in the nmmhc-a gene cause autosomal dominant macrothrombocytopenia with leukocyte inclusions (may-hegglin anomaly/sebastian syndrome). *Blood* **97**(4), 1147–9, Feb (2001).

125. Kelley, M. J., Jawien, W., Ortel, T. L., and Korczak, J. F. Mutation of myh9, encoding non-muscle myosin heavy chain a, in may-hegglin anomaly. *Nat Genet* **26**(1), 106–8, Sep (2000).

126. Sabri, S., Foudi, A., Boukour, S., Franc, B., Charrier, S., Jandrot-Perrus, M., Farndale, R. W., Jalil, A., Blundell, M. P., Cramer, E. M., Louache, F., Debili, N., Thrasher, A. J., and Vainchenker, W. Deficiency in the wiskott-aldrich protein induces premature proplatelet formation and platelet production in the bone marrow compartment. *Blood* **108**(1), 134–40, Jul (2006).

127. Snapper, S. B., Rosen, F. S., Mizoguchi, E., Cohen, P., Khan, W., Liu, C. H., Hagemann, T. L., Kwan, S. P., Ferrini, R., Davidson, L., Bhan, A. K., and Alt, F. W. Wiskott-aldrich syndrome protein-deficient mice reveal a role for wasp in t but not b cell activation. *Immunity* **9**(1), 81–91, Jul (1998).

128. Bender, M., Stritt, S., Nurden, P., van Eeuwijk, J. M. M., Zieger, B., Kentouche, K., Schulze, H., Morbach, H., Stegner, D., Heinze, K. G., Heinze, K., Dütting, S., Gupta, S., Witke, W., Falet, H., Fischer, A., Hartwig, J. H., and Nieswandt, B. Megakaryocyte-specific profilin1-deficiency alters microtubule stability and causes a wiskott-aldrich syndrome-like platelet defect. *Nat Commun* **5**, 4746, Sep (2014).

129. Stritt, S., Birkholz, I., Beck, S., Sorrentino, S., Sapra, K. T., Viaud, J., Heck, J., Gaits-Iacovoni, F., Schulze, H., Du, X., Hartwig, J. H., Braun, A., Bender, M., Medalia, O., and Nieswandt, B. Profilin 1-mediated cytoskeletal rearrangements regulate integrin function in mouse platelets. *Blood Adv* **2**(9), 1040–1045, May (2018).

130. Paul, D. S., Casari, C., Wu, C., Piatt, R., Pasala, S., Campbell, R. A., Poe, K. O., Ghalloussi, D., Lee, R. H., Rotty, J. D., Cooley, B. C., Machlus, K. R., Italiano, Jr, J. E., Weyrich, A. S., Bear, J. E., and Bergmeier, W. Deletion of the arp2/3 complex in megakaryocytes leads to microthrombocytopenia in mice. *Blood Adv* **1**(18), 1398–1408, Aug (2017).

131. Patel-Hett, S., Wang, H., Begonja, A. J., Thon, J. N., Alden, E. C., Wandersee, N. J., An, X., Mohandas, N., Hartwig, J. H., and Italiano, Jr, J. E. The spectrin-based membrane skeleton stabilizes mouse megakaryocyte membrane systems and is essential for proplatelet and platelet formation. *Blood* **118**(6), 1641–52, Aug (2011).

132. Wandersee, N. J., Lee, J. C., Kaysser, T. M., Bronson, R. T., and Barker, J. E. Hematopoietic cells from -spectrin-deficient mice are sufficient to induce thrombotic events in hematopoietically ablated recipients. *Blood* **92**(12), 4856–63, Dec (1998).

133. McArthur, K., Chappaz, S., and Kile, B. T. Apoptosis in megakaryocytes and platelets: the life and death of a lineage. *Blood* **131**(6), 605–610, Feb (2018).

134. Zauli, G., Vitale, M., Falcieri, E., Gibellini, D., Bassini, A., Celeghini, C., Columbaro, M., and Capitani, S. In vitro senescence and apoptotic cell death of human megakaryocytes. *Blood* **90**(6), 2234–43, Sep (1997).

135. De Botton, S., Sabri, S., Daugas, E., Zermati, Y., Guidotti, J. E., Hermine, O., Kroemer, G., Vainchenker, W., and Debili, N. Platelet formation is the consequence of caspase activation within megakaryocytes. *Blood* **100**(4), 1310–7, Aug (2002).

136. Thon, J. N., Montalvo, A., Patel-Hett, S., Devine, M. T., Richardson, J. L., Ehrlicher, A., Larson, M. K., Hoffmeister, K., Hartwig, J. H., and Italiano, Jr, J. E. Cytoskeletal mechanics of proplatelet maturation and platelet release. *J Cell Biol* **191**(4), 861–74, Nov (2010).

137. Thon, J. N., Macleod, H., Begonja, A. J., Zhu, J., Lee, K.-C., Mogilner, A., Hartwig, J. H., and Italiano, Jr, J. E. Microtubule and cortical forces determine platelet size during vascular platelet production. *Nat Commun* **3**, 852, May (2012).

138. Savoia, A. Molecular basis of inherited thrombocytopenias: an update. *Curr Opin Hematol* **23**(5), 486–92, 09 (2016).

139. Johnson, B., Fletcher, S. J., and Morgan, N. V. Inherited thrombocytopenia: novel insights into megakaryocyte maturation, proplatelet formation and platelet lifespan. *Platelets* **27**(6), 519–25, Sep (2016).

140. Johnson, B., Lowe, G. C., Futterer, J., Lordkipanidzé, M., MacDonald, D., Simpson, M. A., Sanchez-Guiú, I., Drake, S., Bem, D., Leo, V., Fletcher, S. J., Dawood, B., Rivera, J., Allsup, D., Biss, T., Bolton-Maggs, P. H., Collins, P., Curry, N., Grimley, C., James, B., Makris, M., Motwani, J., Pavord, S., Talks, K., Thachil, J., Wilde, J., Williams, M., Harrison, P., Gissen, P., Mundell, S., Mumford, A., Daly, M. E., Watson, S. P., Morgan, N. V., and UK GAPP Study Group. Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. *Haematologica* **101**(10), 1170–1179, 10 (2016).

141. Balduini, C. L. and Savoia, A. Genetics of familial forms of thrombocytopenia. *Hum Genet* **131**(12), 1821–32, Dec (2012).

142. Morowski, M., Vögtle, T., Kraft, P., Kleinschnitz, C., Stoll, G., and Nieswandt, B. Only severe thrombocytopenia results in bleeding and defective thrombus formation in mice. *Blood* **121**(24), 4938–47, Jun (2013).

143. Nurden, A. T. Qualitative disorders of platelets and megakaryocytes. *J Thromb Haemost* **3**(8), 1773–82, Aug (2005).

144. Nurden, A. T., Freson, K., and Seligsohn, U. Inherited platelet disorders. *Haemophilia* **18 Suppl 4**, 154–60, Jul (2012).

145. Nurden, A. T. and Nurden, P. Inherited disorders of platelet function: selected updates. *J Thromb Haemost* **13 Suppl 1**, S2–9, Jun (2015).

146. Nurden, A. T. and Nurden, P. Inherited thrombocytopenias. *Haematologica* **92**(9), 1158–64, Sep (2007).

147. Levy-Toledano, S., Caen, J. P., Breton-Gorius, J., Rendu, F., Cywiner-Golenzer, C., Dupuy, E., Legrand, Y., and Maclouf, J. Gray platelet syndrome: alpha-granule deficiency. its influence on platelet function. *J Lab Clin Med* **98**(6), 831–48, Dec (1981).

148. White, J. G. Ultrastructural studies of the gray platelet syndrome. *Am J Pathol* **95**(2), 445–62, May (1979).

149. Gunay-Aygun, M., Zivony-Elboum, Y., Gumruk, F., Geiger, D., Cetin, M., Khayat, M., Kleta, R., Kfir, N., Anikster, Y., Chezar, J., Arcos-Burgos, M., Shalata, A., Stanescu, H., Manaster, J., Arat, M., Edwards, H., Freiberg, A. S., Hart, P. S., Riney, L. C., Patzel, K., Tanpaiboon, P., Markello, T., Huizing, M., Maric, I., Horne, M., Kehrel, B. E., Jurk, K., Hansen, N. F., Cherukuri, P. F., Jones, M., Cruz, P., Mullikin, J. C., Nurden, A., White, J. G., Gahl, W. A., and Falik-Zaccai, T. Gray platelet syndrome: natural history of a large patient cohort and locus assignment to chromosome 3p. *Blood* **116**(23), 4990–5001, Dec (2010).

150. Freson, K., Wijgaerts, A., and van Geet, C. Update on the causes of platelet disorders and functional consequences. *Int J Lab Hematol* **36**(3), 313–25, Jun (2014).

151. Gunay-Aygun, M., Falik-Zaccai, T. C., Vilboux, T., Zivony-Elboum, Y., Gumruk, F., Cetin, M., Khayat, M., Boerkoel, C. F., Kfir, N., Huang, Y., Maynard, D., Dorward, H., Berger, K., Kleta, R., Anikster, Y., Arat, M., Freiberg, A. S., Kehrel, B. E., Jurk, K., Cruz, P., Mullikin, J. C., White, J. G., Huizing,

M., and Gahl, W. A. Nbeal2 is mutated in gray platelet syndrome and is required for biogenesis of platelet -granules. *Nat Genet* **43**(8), 732–4, Jul (2011).

152. Kahr, W. H. A., Lo, R. W., Li, L., Pluthero, F. G., Christensen, H., Ni, R., Vaezzadeh, N., Hawkins, C. E., Weyrich, A. S., Di Paola, J., Landolt-Marticorena, C., and Gross, P. L. Abnormal megakaryocyte development and platelet function in nbeal2(-/-) mice. *Blood* **122**(19), 3349–58, Nov (2013).

153. Boeckelmann, D., Hengartner, H., Greinacher, A., Nowak-Göttl, U., Sachs, U. J., Peter, K., Sandrock-Lang, K., and Zieger, B. Patients with bernard-soulier syndrome and different severity of the bleeding phenotype. *Blood Cells Mol Dis* **67**, 69–74, Sep (2017).

154. Berndt, M. C. and Andrews, R. K. Bernard-soulier syndrome. *Haematologica* **96**(3), 355–9, Mar (2011).

155. Chen, Z. and Shivdasani, R. A. Regulation of platelet biogenesis: insights from the may-hegglin anomaly and other myh9-related disorders. *J Thromb Haemost* **7 Suppl 1**, 272–6, Jul (2009).

156. Chen, Z., Naveiras, O., Balduini, A., Mammoto, A., Conti, M. A., Adelstein, R. S., Ingber, D., Daley, G. Q., and Shivdasani, R. A. The may-hegglin anomaly gene myh9 is a negative regulator of platelet biogenesis modulated by the rho-rock pathway. *Blood* **110**(1), 171–9, Jul (2007).

157. Seri, M., Cusano, R., Gangarossa, S., Caridi, G., Bordo, D., Lo Nigro, C., Ghiggeri, G. M., Ravazzolo, R., Savino, M., Del Vecchio, M., d'Apolito, M., Iolascon, A., Zelante, L. L., Savoia, A., Balduini, C. L., Noris, P., Magrini, U., Belletti, S., Heath, K. E., Babcock, M., Glucksman, M. J., Aliprandis, E., Bizzaro, N., Desnick, R. J., and Martignetti, J. A. Mutations in myh9 result in the may-hegglin anomaly, and fechtner and sebastian syndromes. the may-heggllin/fechtner syndrome consortium. *Nat Genet* **26**(1), 103–5, Sep (2000).

158. Di Paola, J. Paris-trousseau: evidence keeps pointing to fli1. *Blood* **126**(17), 1973–4, Oct (2015).

159. Falet, H. Singling out fli1 in paris-trousseau syndrome. *Blood* **129**(26), 3399–3401, 06 (2017).

160. Ichimiya, Y., Wada, Y., Kunishima, S., Tsukamoto, K., Kosaki, R., Sago, H., Ishiguro, A., and Ito, Y. 11q23 deletion syndrome (jacobsen syndrome) with severe bleeding: a case report. *J Med Case Rep* **12**(1), 3, Jan (2018).

161. Vo, K. K., Jarocha, D. J., Lyde, R. B., Hayes, V., Thom, C. S., Sullivan, S. K., French, D. L., and Poncz, M. Fli1 level during megakaryopoiesis affects thrombopoiesis and platelet biology. *Blood* **129**(26), 3486–3494, 06 (2017).

162. Massaad, M. J., Ramesh, N., and Geha, R. S. Wiskott-aldrich syndrome: a comprehensive review. *Ann N Y Acad Sci* **1285**, 26–43, May (2013).

163. Kahr, W. H. A., Pluthero, F. G., Elkadri, A., Warner, N., Drobac, M., Chen, C. H., Lo, R. W., Li, L., Li, R., Li, Q., Thoeni, C., Pan, J., Leung, G., Lara-Corrales, I., Murchie, R., Cutz, E., Laxer, R. M., Upton, J., Roifman, C. M., Yeung, R. S. M., Brumell, J. H., and Muise, A. M. Loss of the arp2/3 complex component arpc1b causes platelet abnormalities and predisposes to inflammatory disease. *Nat Commun* **8**, 14816, Apr (2017).

164. de Jong, A. and Eikenboom, J. Von willebrand disease mutation spectrum and associated mutation mechanisms. *Thromb Res* **159**, 65–75, Nov (2017).

165. McGinnis, E. and Vercauteren, S. M. von willebrand disease type 2b. *Blood* **128**(23), 2743, 12 (2016).

166. McGinnis, E. and Vercauteren, S. M. von willebrand disease type 2b. *Blood* **129**(4), 538, 01 (2017).

167. Echahdi, H., El Hasbaoui, B., El Khorassani, M., Agadr, A., and Khattab, M. Von willebrand's disease: case report and review of literature. *Pan Afr Med J* **27**, 147 (2017).

168. Nurden, P., Debili, N., Vainchenker, W., Bobe, R., Bredoux, R., Corvazier, E., Combrie, R., Fressinaud, E., Meyer, D., Nurden, A. T., and Enouf, J. Impaired megakaryocytopoiesis in type 2b von willebrand disease with severe thrombocytopenia. *Blood* **108**(8), 2587–95, Oct (2006).

169. Watson, S. P., Lowe, G. C., Lordkipanidzé, M., Morgan, N. V., and GAPP consortium. Genotyping and phenotyping of platelet function disorders. *J Thromb Haemost* **11 Suppl 1**, 351–63, Jun (2013).

170. Lowe, G. C., Lordkipanidzé, M., Watson, S. P., and UK GAPP study group. Utility of the isth bleeding assessment tool in predicting platelet defects in participants with suspected inherited platelet function disorders. *J Thromb Haemost* **11**(9), 1663–8, Sep (2013).

171. Fletcher, S. J., Johnson, B., Lowe, G. C., Bem, D., Drake, S., Lordkipanidzé, M., Guiú, I. S., Dawood, B., Rivera, J., Simpson, M. A., Daly, M. E., Motwani, J., Collins, P. W., Watson, S. P., Morgan, N. V., and UK Genotyping and Phenotyping of Platelets study group. Slfn14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *J Clin Invest* **125**(9), 3600–5, Sep (2015).

172. Huang, B., Babcock, H., and Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **143**(7), 1047–58, Dec (2010).

173. Heintzmann, R. and Huser, T. Super-resolution structured illumination microscopy. *Chem Rev* **117**(23), 13890–13908, Dec (2017).

174. Heintzmann, R. Saturated patterned excitation microscopy with two-dimensional excitation patterns. *Micron* **34**(6-7), 283–91 (2003).

175. Gustafsson, M. G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J Microsc* **198**(Pt 2), 82–7, May (2000).

176. Hirvonen, L. M., Wicker, K., Mandula, O., and Heintzmann, R. Structured illumination microscopy of a living cell. *Eur Biophys J* **38**(6), 807–12, Jul (2009).

177. Gliko, O., Brownell, W. E., and Saggau, P. Fast two-dimensional standing-wave total-internal-reflection fluorescence microscopy using acousto-optic deflectors. *Opt Lett* **34**(6), 836–8, Mar (2009).

178. Hoyer, P., de Medeiros, G., Balázs, B., Norlin, N., Besir, C., Hanne, J., Kräusslich, H.-G., Engelhardt, J., Sahl, S. J., Hell, S. W., and Hufnagel, L. Breaking the diffraction limit of light-sheet fluorescence microscopy by resolft. *Proc Natl Acad Sci U S A* **113**(13), 3442–6, Mar (2016).

179. Grotjohann, T., Testa, I., Leutenegger, M., Bock, H., Urban, N. T., Lavoie-Cardinal, F., Willig, K. I., Eggeling, C., Jakobs, S., and Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic gfp. *Nature* **478**(7368), 204–8, Sep (2011).

180. Chang, B.-J., Perez Meza, V. D., and Stelzer, E. H. K. csilsfm combines light-sheet fluorescence microscopy and coherent structured illumination for a lateral resolution below 100 nm. *Proc Natl Acad Sci U S A* **114**(19), 4869–4874, 05 (2017).

181. Westmoreland, D., Shaw, M., Grimes, W., Metcalf, D. J., Burden, J. J., Gomez, K., Knight, A. E., and Cutler, D. F. Super-resolution microscopy as a potential approach to diagnosis of platelet granule disorders. *J Thromb Haemost* **14**(4), 839–49, Apr (2016).

182. Blom, H. and Widengren, J. Stimulated emission depletion microscopy. *Chem Rev* **117**(11), 7377–7427, Jun (2017).

183. Klar, T. A., Engel, E., and Hell, S. W. Breaking abbe's diffraction resolution limit in fluorescence microscopy with stimulated emission depletion beams of various shapes. *Phys Rev E Stat Nonlin Soft Matter Phys* **64**(6 Pt 2), 066613, Dec (2001).

184. Hell, S. W. and Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt Lett* **19**(11), 780–2, Jun (1994).

185. Roobala, C., Ilanila, I. P., and Basu, J. K. Applications of sted fluorescence nanoscopy in unravelling nanoscale structure and dynamics of biological systems. *J Biosci* **43**(3), 471–484, Jul (2018).

186. Gao, M., Maraspini, R., Beutel, O., Zehtabian, A., Eickholt, B., Honigmann, A., and Ewers, H. Expansion stimulated emission depletion microscopy (exsted). *ACS Nano* **12**(5), 4178–4185, May (2018).

187. Dertinger, T., Colyer, R., Iyer, G., Weiss, S., and Enderlein, J. Fast, background-free, 3d super-resolution optical fluctuation imaging (sofi). *Proc Natl Acad Sci U S A* **106**(52), 22287–92, Dec (2009).

188. Cox, S., Rosten, E., Monypenny, J., Jovanovic-Talisman, T., Burnette, D. T., Lippincott-Schwartz, J., Jones, G. E., and Heintzmann, R. Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nat Methods* **9**(2), 195–200, Dec (2011).

189. Gustafsson, N., Culley, S., Ashdown, G., Owen, D. M., Pereira, P. M., and Henriques, R. Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations. *Nat Commun* **7**, 12471, Aug (2016).

190. Huisken, J. and Stainier, D. Y. R. Selective plane illumination microscopy techniques in developmental biology. *Development* **136**(12), 1963–75, Jun (2009).

191. Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E. H. K. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* **305**(5686), 1007–9, Aug (2004).

192. Sahl, S. J., Balzarotti, F., Keller-Findeisen, J., Leutenegger, M., Westphal, V., Egner, A., Lavoie-Cardinal, F., Chmyrov, A., Grotjohann, T., and Jakobs, S. Comment on "extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics". *Science* **352**(6285), 527, Apr (2016).

193. Li, D., Shao, L., Chen, B.-C., Zhang, X., Zhang, M., Moses, B., Milkie, D. E., Beach, J. R., Hammer, 3rd, J. A., Pasham, M., Kirchhausen, T., Baird, M. A., Davidson, M. W., Xu, P., and Betzig, E. Advanced imaging. extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* **349**(6251), aab3500, Aug (2015).

194. Gustavsson, A.-K., Petrov, P. N., Lee, M. Y., Shechtman, Y., and Moerner, W. E. 3d single-molecule super-resolution microscopy with a tilted light sheet. *Nat Commun* **9**(1), 123, Jan (2018).

195. Hu, Y. S., Zimmerley, M., Li, Y., Watters, R., and Cang, H. Single-molecule super-resolution light-sheet microscopy. *Chemphyschem* **15**(4), 577–86, Mar (2014).

196. Mir, M., Reimer, A., Stadler, M., Tangara, A., Hansen, A. S., Hockemeyer, D., Eisen, M. B., Garcia, H., and Darzacq, X. Single molecule imaging in live embryos using lattice light-sheet microscopy. *Methods Mol Biol* **1814**, 541–559 (2018).

197. Bates, M., Huang, B., Dempsey, G. T., and Zhuang, X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* **317**(5845), 1749–53, Sep (2007).

198. Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**(5793), 1642–5, Sep (2006).

199. Rust, M. J., Bates, M., and Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nat Methods* **3**(10), 793–5, Oct (2006).

200. Qu, X., Wu, D., Mets, L., and Scherer, N. F. Nanometer-localized multiple single-molecule fluorescence microscopy. *Proc Natl Acad Sci U S A* **101**(31), 11298–303, Aug (2004).

201. Miklosi, A. G., Del Favero, G., Marko, D., Harkany, T., and Lubec, G. Resolution matters: Correlating quantitative proteomics and nanoscale-precision microscopy for reconstructing synapse identity. *Proteomics* **18**(14), e1800139, Jul (2018).

202. Mayr, S., Hauser, F., Peterbauer, A., Tauscher, A., Naderer, C., Axmann, M., Plochberger, B., and Jacak, J. Localization microscopy of actin cytoskeleton in human platelets. *Int J Mol Sci* **19**(4), Apr (2018).

203. Sengupta, P. and Lippincott-Schwartz, J. Quantitative analysis of photoactivated localization microscopy (palm) datasets using pair-correlation analysis. *Bioessays* **34**(5), 396–405, May (2012).

204. Sengupta, P., Jovanovic-Talisman, T., Skoko, D., Renz, M., Veatch, S. L., and Lippincott-Schwartz, J. Probing protein heterogeneity in the plasma membrane using palm and pair correlation analysis. *Nat Methods* **8**(11), 969–75, Sep (2011).

205. Burnette, D. T., Sengupta, P., Dai, Y., Lippincott-Schwartz, J., and Kachar, B. Bleaching/blinking assisted localization microscopy for superresolution imaging using standard fluorescent molecules. *Proc Natl Acad Sci U S A* **108**(52), 21081–6, Dec (2011).

206. Ratz, M., Testa, I., Hell, S. W., and Jakobs, S. Crispr/cas9-mediated endogenous protein tagging for resolft super-resolution microscopy of living human cells. *Sci Rep* **5**, 9592, Apr (2015).

207. Cho, W.-K., Jayanth, N., Mullen, S., Tan, T. H., Jung, Y. J., and Cissé, I. I. Super-resolution imaging of fluorescently labeled, endogenous rna polymerase ii in living cells with crispr/cas9-mediated gene editing. *Sci Rep* **6**, 35949, 10 (2016).

208. Ovesný, M., Křížek, P., Borkovec, J., Svindrych, Z., and Hagen, G. M. Thunderstorm: a comprehensive imagej plug-in for palm and storm data analysis and super-resolution imaging. *Bioinformatics* **30**(16), 2389–90, Aug (2014).

209. Culley, S., Albrecht, D., Jacobs, C., Pereira, P. M., Leterrier, C., Mercer, J., and Henriques, R. Quantitative mapping and minimization of super-resolution optical imaging artifacts. *Nat Methods* **15**(4), 263–266, Apr (2018).

210. Banterle, N., Bui, K. H., Lemke, E. A., and Beck, M. Fourier ring correlation as a resolution criterion for super-resolution microscopy. *J Struct Biol* **183**(3), 363–367, Sep (2013).

211. Habuchi, S., Dedecker, P., Hotta, J.-i., Flors, C., Ando, R., Mizuno, H., Miyawaki, A., and Hofkens, J. Photo-induced protonation/deprotonation in the gfp-like fluorescent protein dronpa: mechanism responsible for the reversible photoswitching. *Photochem Photobiol Sci* **5**(6), 567–76, Jun (2006).

212. Habuchi, S., Ando, R., Dedecker, P., Verheijen, W., Mizuno, H., Miyawaki, A., and Hofkens, J. Reversible single-molecule photoswitching in the gfp-like fluorescent protein dronpa. *Proc Natl Acad Sci U S A* **102**(27), 9511–6, Jul (2005).

213. McKinney, S. A., Murphy, C. S., Hazelwood, K. L., Davidson, M. W., and Looger, L. L. A bright and photostable photoconvertible fluorescent protein. *Nat Methods* **6**(2), 131–3, Feb (2009).

214. Zhang, M., Chang, H., Zhang, Y., Yu, J., Wu, L., Ji, W., Chen, J., Liu, B., Lu, J., Liu, Y., Zhang, J., Xu, P., and Xu, T. Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat Methods* **9**(7), 727–9, May (2012).

215. Nieuwenhuizen, R. P. J., Lidke, K. A., Bates, M., Puig, D. L., Grünwald, D., Stallinga, S., and Rieger, B. Measuring image resolution in optical nanoscopy. *Nat Methods* **10**(6), 557–62, Jun (2013).

216. Paez-Segala, M. G., Sun, M. G., Shtengel, G., Viswanathan, S., Baird, M. A., Macklin, J. J., Patel, R., Allen, J. R., Howe, E. S., Piszczek, G., Hess, H. F., Davidson, M. W., Wang, Y., and Looger, L. L. Fixation-resistant photoactivatable fluorescent proteins for clem. *Nat Methods* **12**(3), 215–8, 4 p following 218, Mar (2015).

217. Los, G. V. and Wood, K. The halotag: a novel technology for cell imaging and protein analysis. *Methods Mol Biol* **356**, 195–208 (2007).

218. Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G.,

Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., and Wood, K. V. Halotag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **3**(6), 373–82, Jun (2008).

219. Roberts, B., Haupt, A., Tucker, A., Grancharova, T., Arakaki, J., Fuqua, M. A., Nelson, A., Hookway, C., Ludmann, S. A., Mueller, I. A., Yang, R., Horwitz, R., Rafelski, S. M., and Gunawardane, R. N. Systematic gene tagging using crispr/cas9 in human stem cells to illuminate cell organization. *Mol Biol Cell* **28**(21), 2854–2874, Oct (2017).

220. Cranfill, P. J., Sell, B. R., Baird, M. A., Allen, J. R., Lavagnino, Z., de Gruiter, H. M., Kremers, G.-J., Davidson, M. W., Ustione, A., and Piston, D. W. Quantitative assessment of fluorescent proteins. *Nat Methods* **13**(7), 557–62, 07 (2016).

221. Uemura, T., Mori, T., Kurihara, T., Kawase, S., Koike, R., Satoga, M., Cao, X., Li, X., Yanagawa, T., Sakurai, T., Shindo, T., and Tabuchi, K. Fluorescent protein tagging of endogenous protein in brain neurons using crispr/cas9-mediated knock-in and in utero electroporation techniques. *Sci Rep* **6**, 35861, 10 (2016).

222. Komatsu, N. Culture of megakaryocytic cell lines: uses and limitations. *Methods Mol Biol* **272**, 361–73 (2004).

223. Schulze, H. Culture of murine megakaryocytes and platelets from fetal liver and bone marrow. *Methods Mol Biol* **788**, 193–203 (2012).

224. Schulze, H. Culture, expansion, and differentiation of murine megakaryocytes from fetal liver, bone marrow, and spleen. *Curr Protoc Immunol* **112**, 22F.6.1–22F.6.15, Feb (2016).

225. Balduini, A., Pallotta, I., Malara, A., Lova, P., Pecci, A., Viarengo, G., Balduini, C. L., and Torti, M. Adhesive receptors, extracellular proteins and myosin iia orchestrate proplatelet formation by human megakaryocytes. *J Thromb Haemost* **6**(11), 1900–7, Nov (2008).

226. Schuhmann, N. K., Pozzoli, O., Sallach, J., Huber, A., Avitabile, D., Perabo, L., Rappl, G., Capogrossi, M. C., Hallek, M., Pesce, M., and Büning, H. Gene transfer into human cord blood-derived cd34(+) cells by adeno-associated viral vectors. *Exp Hematol* **38**(9), 707–17, Sep (2010).

227. Takahashi, K. and Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**(4), 663–76, Aug (2006).

228. Gurdon, J. B. Adult frogs derived from the nuclei of single somatic cells. *Dev Biol* **4**, 256–73, Apr (1962).

229. Gurdon, J. B. The transplantation of nuclei between two species of xenopus. *Dev Biol* **5**, 68–83, Aug (1962).

230. Campbell, K. H., McWhir, J., Ritchie, W. A., and Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380**(6569), 64–6, Mar (1996).

231. Avior, Y., Sagi, I., and Benvenisty, N. Pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Mol Cell Biol* **17**(3), 170–82, Mar (2016).

232. Borst, S., Sim, X., Poncz, M., French, D. L., and Gadue, P. Induced pluripotent stem cell-derived megakaryocytes and platelets for disease modeling and future clinical applications. *Arterioscler Thromb Vasc Biol* **37**(11), 2007–2013, 11 (2017).

233. Thon, J. N., Mazutis, L., Wu, S., Sylman, J. L., Ehrlicher, A., Machlus, K. R., Feng, Q., Lu, S., Lanza, R., Neeves, K. B., Weitz, D. A., and Italiano, Jr, J. E. Platelet bioreactor-on-a-chip. *Blood* **124**(12), 1857–67, Sep (2014).

234. Murry, C. E. and Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**(4), 661–80, Feb (2008).

235. Lu, S.-J., Li, F., Yin, H., Feng, Q., Kimbrel, E. A., Hahm, E., Thon, J. N., Wang, W., Italiano, J. E., Cho, J., and Lanza, R. Platelets generated from human embryonic stem cells are functional in vitro and in the microcirculation of living mice. *Cell Res* **21**(3), 530–45, Mar (2011).

236. Lu, S.-J., Feng, Q., Park, J. S., and Lanza, R. Directed differentiation of red blood cells from human embryonic stem cells. *Methods Mol Biol* **636**, 105–21 (2010).

237. Lu, S.-J., Feng, Q., Caballero, S., Chen, Y., Moore, M. A. S., Grant, M. B., and Lanza, R. Generation of functional hemangioblasts from human embryonic stem cells. *Nat Methods* **4**(6), 501–9, Jun (2007).

238. Feng, Q., Shabrani, N., Thon, J. N., Huo, H., Thiel, A., Machlus, K. R., Kim, K., Brooks, J., Li, F., Luo, C., Kimbrel, E. A., Wang, J., Kim, K.-S., Italiano, J., Cho, J., Lu, S.-J., and Lanza, R. Scalable generation of universal platelets from human induced pluripotent stem cells. *Stem Cell Reports* **3**(5), 817–31, Nov (2014).

239. Borst, S., Sim, X., Poncz, M., French, D. L., and Gadue, P. Induced pluripotent stem cell-derived megakaryocytes and platelets for disease modeling and future clinical applications. *Arterioscler Thromb Vasc Biol* **37**(11), 2007–2013, Nov (2017).

240. Nakamura, S., Takayama, N., Hirata, S., Seo, H., Endo, H., Ochi, K., Fujita, K.-i., Koike, T., Harimoto, K.-i., Dohda, T., Watanabe, A., Okita, K., Takahashi, N., Sawaguchi, A., Yamanaka, S., Nakauchi, H., Nishimura, S., and Eto, K. Expandable megakaryocyte cell lines enable clinically applicable generation of platelets from human induced pluripotent stem cells. *Cell Stem Cell* **14**(4), 535–48, Apr (2014).

241. Thon, J. N., Dykstra, B. J., and Beaulieu, L. M. Platelet bioreactor: accelerated evolution of design and manufacture. *Platelets* **28**(5), 472–477, Jan (2017).

242. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity. *Science* **337**(6096), 816–21, Aug (2012).

243. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. Multiplex genome engineering using crispr/cas systems. *Science* **339**(6121), 819–23, Feb (2013).

244. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. Rna-guided human genome engineering via cas9. *Science* **339**(6121), 823–6, Feb (2013).

245. Razzouk, S. Crispr-cas9: A cornerstone for the evolution of precision medicine. *Ann Hum Genet* **Epub**, Jul (2018).

246. Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. Double nicking by rna-guided crispr cas9 for enhanced genome editing specificity. *Cell* **154**(6), 1380–9, Sep (2013).

247. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. Genome engineering using the crispr-cas9 system. *Nat Protoc* **8**(11), 2281–2308, Nov (2013).

248. Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., and Qi, L. S. Crispr-mediated modular rna-guided regulation of transcription in eukaryotes. *Cell* **154**(2), 442–51, Jul (2013).

249. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I. T., and Yamamoto, T. Mmej-assisted gene knock-in using talens and crispr-cas9 with the pitch systems. *Nat Protoc* **11**(1), 118–33, Jan (2016).

250. Iizuka, H., Kagoya, Y., Kataoka, K., Yoshimi, A., Miyauchi, M., Taoka, K., Kumano, K., Yamamoto, T., Hotta, A., Arai, S., and Kurokawa, M. Targeted gene correction of runx1 in induced pluripotent

stem cells derived from familial platelet disorder with propensity to myeloid malignancy restores normal megakaryopoiesis. *Exp Hematol* **43**(10), 849–57, Oct (2015).

251. Connelly, J. P., Kwon, E. M., Gao, Y., Trivedi, N. S., Elkahloun, A. G., Horwitz, M. S., Cheng, L., and Liu, P. P. Targeted correction of runx1 mutation in fpd patient-specific induced pluripotent stem cells rescues megakaryopoietic defects. *Blood* **124**(12), 1926–30, Sep (2014).

252. Sakurai, M., Kunimoto, H., Watanabe, N., Fukuchi, Y., Yuasa, S., Yamazaki, S., Nishimura, T., Sadahira, K., Fukuda, K., Okano, H., Nakauchi, H., Morita, Y., Matsumura, I., Kudo, K., Ito, E., Ebihara, Y., Tsuji, K., Harada, Y., Harada, H., Okamoto, S., and Nakajima, H. Impaired hematopoietic differentiation of runx1-mutated induced pluripotent stem cells derived from fpd/aml patients. *Leukemia* **28**(12), 2344–54, Dec (2014).

253. Sullivan, S. K., Mills, J. A., Koukouritaki, S. B., Vo, K. K., Lyde, R. B., Paluru, P., Zhao, G., Zhai, L., Sullivan, L. M., Wang, Y., Kishore, S., Gharaibeh, E. Z., Lambert, M. P., Wilcox, D. A., French, D. L., Poncz, M., and Gadue, P. High-level transgene expression in induced pluripotent stem cell-derived megakaryocytes: correction of glanzmann thrombasthenia. *Blood* **123**(5), 753–7, Jan (2014).

254. Jacobi, A. M., Rettig, G. R., Turk, R., Collingwood, M. A., Zeiner, S. A., Quadros, R. M., Harms, D. W., Bonthuis, P. J., Gregg, C., Ohtsuka, M., Gurumurthy, C. B., and Behlke, M. A. Simplified crispr tools for efficient genome editing and streamlined protocols for their delivery into mammalian cells and mouse zygotes. *Methods* **121-122**, 16–28, 05 (2017).

255. Li, G., Zhang, X., Zhong, C., Mo, J., Quan, R., Yang, J., Liu, D., Li, Z., Yang, H., and Wu, Z. Small molecules enhance crispr/cas9-mediated homology-directed genome editing in primary cells. *Sci Rep* **7**(1), 8943, Aug (2017).

256. Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., and Ploegh, H. L. Increasing the efficiency of precise genome editing with crispr-cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**(5), 538–42, May (2015).

257. Hu, Z., Shi, Z., Guo, X., Jiang, B., Wang, G., Luo, D., Chen, Y., and Zhu, Y.-S. Ligase iv inhibitor scr7 enhances gene editing directed by crispr-cas9 and ssodn in human cancer cells. *Cell Biosci* **8**, 12 (2018).

258. Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. Rs-1 enhances crispr/cas9- and talen-mediated knock-in efficiency. *Nat Commun* **7**, 10548, Jan (2016).

259. Jayathilaka, K., Sheridan, S. D., Bold, T. D., Bochenska, K., Logan, H. L., Weichselbaum, R. R., Bishop, D. K., and Connell, P. P. A chemical compound that stimulates the human homologous recombination protein rad51. *Proc Natl Acad Sci U S A* **105**(41), 15848–53, Oct (2008).

260. Riesenberg, S. and Maricic, T. Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat Commun* **9**(1), 2164, Jun (2018).

261. Lessard, S., Francioli, L., Alfoldi, J., Tardif, J.-C., Ellinor, P. T., MacArthur, D. G., Lettre, G., Orkin, S. H., and Canver, M. C. Human genetic variation alters crispr-cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc Natl Acad Sci U S A* **114**(52), E11257–E11266, 12 (2017).

262. Bar-Nur, O., Russ, H. A., Efrat, S., and Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* **9**(1), 17–23, Jul (2011).

263. Gadadhar, S., Dadi, H., Bodakuntla, S., Schnitzler, A., Bièche, I., Rusconi, F., and Janke, C. Tubulin glycylation controls primary cilia length. *J Cell Biol* **216**(9), 2701–2713, 09 (2017).

264. Gadadhar, S., Bodakuntla, S., Natarajan, K., and Janke, C. The tubulin code at a glance. *J Cell Sci* **130**(8), 1347–1353, 04 (2017).

265. Gupta, A. and Kitagawa, D. Ultrastructural diversity between centrioles of eukaryotes. *J Biochem* **164**(1), 1–8, Jul (2018).

266. Ludueña, R. F. A hypothesis on the origin and evolution of tubulin. *Int Rev Cell Mol Biol* **302**, 41–185 (2013).

267. Maruta, H., Greer, K., and Rosenbaum, J. L. The acetylation of alpha-tubulin and its relationship to the assembly and disassembly of microtubules. *J Cell Biol* **103**(2), 571–9, Aug (1986).

268. Paturle-Lafanechère, L., Eddé, B., Denoulet, P., Van Dorsselaer, A., Mazarguil, H., Le Caer, J. P., Wehland, J., and Job, D. Characterization of a major brain tubulin variant which cannot be tyrosinated. *Biochemistry* **30**(43), 10523–8, Oct (1991).

269. Aillaud, C., Bosc, C., Saoudi, Y., Denarier, E., Peris, L., Sago, L., Taulet, N., Cieren, A., Tort, O., Magiera, M. M., Janke, C., Redeker, V., Andrieux, A., and Moutin, M.-J. Evidence for new c-terminally truncated variants of - and -tubulins. *Mol Biol Cell* **27**(4), 640–53, Feb (2016).

270. Khawaja, S., Gundersen, G. G., and Bulinski, J. C. Enhanced stability of microtubules enriched in detyrosinated tubulin is not a direct function of detyrosination level. *J Cell Biol* **106**(1), 141–9, Jan (1988).

271. Peris, L., Wagenbach, M., Lafanechère, L., Brocard, J., Moore, A. T., Kozielski, F., Job, D., Wordeman, L., and Andrieux, A. Motor-dependent microtubule disassembly driven by tubulin tyrosination. *J Cell Biol* **185**(7), 1159–66, Jun (2009).

272. Liao, G. and Gundersen, G. G. Kinesin is a candidate for cross-bridging microtubules and intermediate filaments. selective binding of kinesin to detyrosinated tubulin and vimentin. *J Biol Chem* **273**(16), 9797–803, Apr (1998).

273. Magiera, M. M., Singh, P., and Janke, C. Snapshot: Functions of tubulin posttranslational modifications. *Cell* **173**(6), 1552–1552.e1, May (2018).

274. Prota, A. E., Magiera, M. M., Kuijpers, M., Bargsten, K., Frey, D., Wieser, M., Jaussi, R., Hoogenraad, C. C., Kammerer, R. A., Janke, C., and Steinmetz, M. O. Structural basis of tubulin tyrosination by tubulin tyrosine ligase. *J Cell Biol* **200**(3), 259–70, Feb (2013).

275. Kerr, J. P., Robison, P., Shi, G., Bogush, A. I., Kempema, A. M., Hexum, J. K., Becerra, N., Harki, D. A., Martin, S. S., Raiteri, R., Prosser, B. L., and Ward, C. W. Detyrosinated microtubules modulate mechanotransduction in heart and skeletal muscle. *Nat Commun* **6**, 8526, Oct (2015).

276. Robison, P., Caporizzo, M. A., Ahmadzadeh, H., Bogush, A. I., Chen, C. Y., Margulies, K. B., Shenoy, V. B., and Prosser, B. L. Detyrosinated microtubules buckle and bear load in contracting cardiomyocytes. *Science* **352**(6284), aaf0659, Apr (2016).

277. Greer, K., Maruta, H., L'Hernault, S. W., and Rosenbaum, J. L. Alpha-tubulin acetylase activity in isolated chlamydomonas flagella. *J Cell Biol* **101**(6), 2081–4, Dec (1985).

278. North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M., and Verdin, E. The human sir2 ortholog, sirt2, is an nad+-dependent tubulin deacetylase. *Mol Cell* **11**(2), 437–44, Feb (2003).

279. Liu, N., Xiong, Y., Li, S., Ren, Y., He, Q., Gao, S., Zhou, J., and Shui, W. New hdac6-mediated deacetylation sites of tubulin in the mouse brain identified by quantitative mass spectrometry. *Sci Rep* **5**, 16869, Nov (2015).

280. Dompierre, J. P., Godin, J. D., Charrin, B. C., Cordelières, F. P., King, S. J., Humbert, S., and Saudou, F. Histone deacetylase 6 inhibition compensates for the transport deficit in huntington's disease by increasing tubulin acetylation. *J Neurosci* **27**(13), 3571–83, Mar (2007).

281. Benoy, V., Van Helleputte, L., Prior, R., d'Ydewalle, C., Haeck, W., Geens, N., Scheveneels, W., Schevenels, B., Cader, M. Z., Talbot, K., Kozikowski, A. P., Vanden Berghe, P., Van Damme, P.,

Robberecht, W., and Van Den Bosch, L. Hdac6 is a therapeutic target in mutant gars-induced charcot-marie-tooth disease. *Brain* **141**(3), 673–687, Mar (2018).

282. d'Ydewalle, C., Krishnan, J., Chiheb, D. M., Van Damme, P., Irobi, J., Kozikowski, A. P., Vanden Berghe, P., Timmerman, V., Robberecht, W., and Van Den Bosch, L. Hdac6 inhibitors reverse axonal loss in a mouse model of mutant hspb1-induced charcot-marie-tooth disease. *Nat Med* **17**(8), 968–74, Jul (2011).

283. Wolff, A., de Néchaud, B., Chillet, D., Mazarguil, H., Desbruyères, E., Audebert, S., Eddé, B., Gros, F., and Denoulet, P. Distribution of glutamylated alpha and beta-tubulin in mouse tissues using a specific monoclonal antibody, gt335. *Eur J Cell Biol* **59**(2), 425–32, Dec (1992).

284. Wloga, D., Webster, D. M., Rogowski, K., Bré, M.-H., Levilliers, N., Jerka-Dziadosz, M., Janke, C., Dougan, S. T., and Gaertig, J. Ttll3 is a tubulin glycine ligase that regulates the assembly of cilia. *Dev Cell* **16**(6), 867–76, Jun (2009).

285. Rogowski, K., Juge, F., van Dijk, J., Wloga, D., Strub, J.-M., Levilliers, N., Thomas, D., Bré, M.-H., Van Dorsselaer, A., Gaertig, J., and Janke, C. Evolutionary divergence of enzymatic mechanisms for posttranslational polyglycylation. *Cell* **137**(6), 1076–87, Jun (2009).

286. Reiter, J. F. and Leroux, M. R. Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol* **18**(9), 533–547, Sep (2017).

287. Konno, A., Ikegami, K., Konishi, Y., Yang, H.-J., Abe, M., Yamazaki, M., Sakimura, K., Yao, I., Shiba, K., Inaba, K., and Setou, M. Ttll9-/- mice sperm flagella show shortening of doublet 7, reduction of doublet 5 polyglutamylation and a stall in beating. *J Cell Sci* **129**(14), 2757–66, 07 (2016).

288. Mullen, R. J., Eicher, E. M., and Sidman, R. L. Purkinje cell degeneration, a new neurological mutation in the mouse. *Proc Natl Acad Sci U S A* **73**(1), 208–12, Jan (1976).

289. Vogel, P., Hansen, G., Fontenot, G., and Read, R. Tubulin tyrosine ligase-like 1 deficiency results in chronic rhinosinusitis and abnormal development of spermatid flagella in mice. *Vet Pathol* **47**(4), 703–12, Jul (2010).

290. Ikegami, K., Sato, S., Nakamura, K., Ostrowski, L. E., and Setou, M. Tubulin polyglutamylation is essential for airway ciliary function through the regulation of beating asymmetry. *Proc Natl Acad Sci U S A* **107**(23), 10490–5, Jun (2010).

291. Marchena, M., Lara, J., Aijón, J., Germain, F., de la Villa, P., and Velasco, A. The retina of the pcd/pcd mouse as a model of photoreceptor degeneration. a structural and functional study. *Exp Eye Res* **93**(5), 607–17, Nov (2011).

292. Astuti, G. D. N., Arno, G., Hull, S., Pierrache, L., Venselaar, H., Carss, K., Raymond, F. L., Collin, R. W. J., Faradz, S. M. H., van den Born, L. I., Webster, A. R., and Cremers, F. P. M. Mutations in agbl5, encoding -tubulin deglutamylase, are associated with autosomal recessive retinitis pigmentosa. *Invest Ophthalmol Vis Sci* **57**(14), 6180–6187, Nov (2016).

293. Branham, S. E., Wright, S. J., Reba, A., and Linder, C. R. Genome-wide association study of arabidopsis thaliana identifies determinants of natural variation in seed oil composition. *J Hered* **107**(3), 248–56, 05 (2016).

294. Bosch Grau, M., Masson, C., Gadadhar, S., Rocha, C., Tort, O., Marques Sousa, P., Vacher, S., Bieche, I., and Janke, C. Alterations in the balance of tubulin glycylation and glutamylation in photoreceptors leads to retinal degeneration. *J Cell Sci* **130**(5), 938–949, 03 (2017).

295. Bosch Grau, M., Masson, C., Gadadhar, S., Rocha, C., Tort, O., Marques Sousa, P., Vacher, S., Bieche, I., and Janke, C. Alterations in the balance of tubulin glycylation and glutamylation in photoreceptors leads to retinal degeneration. *J Cell Sci* **130**(5), 938–949, Mar (2017).

296. Rocha, C., Papon, L., Cacheux, W., Marques Sousa, P., Lascano, V., Tort, O., Giordano, T., Vacher, S., Lemmers, B., Mariani, P., Meseure, D., Medema, J. P., Bièche, I., Hahne, M., and Janke, C. Tubulin glycylases are required for primary cilia, control of cell proliferation and tumor development in colon. *EMBO J* **33**(19), 2247–60, Oct (2014).

297. Ikegami, K., Heier, R. L., Taruishi, M., Takagi, H., Mukai, M., Shimma, S., Taira, S., Hatanaka, K., Morone, N., Yao, I., Campbell, P. K., Yuasa, S., Janke, C., Macgregor, G. R., and Setou, M. Loss of alpha-tubulin polyglutamylation in rosa22 mice is associated with abnormal targeting of kif1a and modulated synaptic function. *Proc Natl Acad Sci U S A* **104**(9), 3213–8, Feb (2007).

298. Janke, C. and Bulinski, J. C. Post-translational regulation of the microtubule cytoskeleton: mechanisms and functions. *Nat Rev Mol Cell Biol* **12**(12), 773–86, Nov (2011).

299. Reck-Peterson, S. L., Redwine, W. B., Vale, R. D., and Carter, A. P. The cytoplasmic dynein transport machinery and its many cargoes. *Nat Rev Mol Cell Biol* **19**(6), 382–398, Jun (2018).

300. Hirokawa, N. and Tanaka, Y. Kinesin superfamily proteins (kifs): Various functions and their relevance for important phenomena in life and diseases. *Exp Cell Res* **334**(1), 16–25, May (2015).

301. Goodson, H. V. and Jonasson, E. M. Microtubules and microtubule-associated proteins. *Cold Spring Harb Perspect Biol* **10**(6), Jun (2018).

302. Bramblett, G. T., Goedert, M., Jakes, R., Merrick, S. E., Trojanowski, J. Q., and Lee, V. M. Abnormal tau phosphorylation at ser396 in alzheimer's disease recapitulates development and contributes to reduced microtubule binding. *Neuron* **10**(6), 1089–99, Jun (1993).

303. Burley, K., Westbury, S. K., and Mumford, A. D. Tubb1 variants and human platelet traits. *Platelets* **29**(2), 209–211, 03 (2018).

304. Dmitrieff, S., Alsina, A., Mathur, A., and Nédélec, F. J. Balance of microtubule stiffness and cortical tension determines the size of blood cells with marginal band across species. *Proc Natl Acad Sci U S A* **114**(17), 4418–4423, Apr (2017).

305. Diagouraga, B., Grichine, A., Fertin, A., Wang, J., Khochbin, S., and Sadoul, K. Motor-driven marginal band coiling promotes cell shape change during platelet activation. *J Cell Biol* **204**(2), 177–85, Jan (2014).

306. Behnke, O. Microtubules in disk-shaped blood cells. *Int Rev Exp Pathol* **9**, 1–92 (1970).

307. Patel-Hett, S., Richardson, J. L., Schulze, H., Drabek, K., Isaac, N. A., Hoffmeister, K., Shivdasani, R. A., Bulinski, J. C., Galjart, N., Hartwig, J. H., and Italiano, Jr, J. E. Visualization of microtubule growth in living platelets reveals a dynamic marginal band with multiple microtubules. *Blood* **111**(9), 4605–16, May (2008).

308. Sadoul, K. New explanations for old observations: marginal band coiling during platelet activation. *J Thromb Haemost* **13**(3), 333–46, Mar (2015).

309. Kunishima, S., Nishimura, S., Suzuki, H., Imaizumi, M., and Saito, H. Tubb1 mutation disrupting microtubule assembly impairs proplatelet formation and results in congenital macrothrombocytopenia. *Eur J Haematol* **92**(4), 276–82, Apr (2014).

310. Fiore, M., Goulas, C., and Pillois, X. A new mutation in tubb1 associated with thrombocytopenia confirms that c-terminal part of 1-tubulin plays a role in microtubule assembly. *Clin Genet* **91**(6), 924–926, Jun (2017).

311. Auer, P. L., Teumer, A., Schick, U., O'Shaughnessy, A., Lo, K. S., Chami, N., Carlson, C., de Denus, S., Dubé, M.-P., Haessler, J., Jackson, R. D., Kooperberg, C., Perreault, L.-P. L., Nauck, M., Peters, U., Rioux, J. D., Schmidt, F., Turcot, V., Völker, U., Völzke, H., Greinacher, A., Hsu, L., Tardif, J.-C., Diaz, G. A., Reiner, A. P., and Lettre, G. Rare and low-frequency coding variants in cxcr2 and other genes are associated with hematological traits. *Nat Genet* **46**(6), 629–34, Jun (2014).

312. Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M. A., Lambourne, J. J., Sivapalaratnam, S., Downes, K., Kundu, K., Bomba, L., Berentsen, K., Bradley, J. R., Daugherty, L. C., Delaneau, O., Freson, K., Garner, S. F., Grassi, L., Guerrero, J., Haimel, M., Janssen-Megens, E. M., Kaan, A., Kamat, M., Kim, B., Mandoli, A., Marchini, J., Martens, J. H. A., Meacham, S., Megy, K., O'Connell, J., Petersen, R., Sharifi, N., Sheard, S. M., Staley, J. R., Tuna, S., van der Ent, M., Walter, K., Wang, S.-Y., Wheeler, E., Wilder, S. P., Iotchkova, V., Moore, C., Sambrook, J., Stunnenberg, H. G., Di Angelantonio, E., Kaptoge, S., Kuijpers, T. W., Carrillo-de Santa-Pau, E., Juan, D., Rico, D., Valencia, A., Chen, L., Ge, B., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yang, Y., Guigo, R., Beck, S., Paul, D. S., Pastinen, T., Bujold, D., Bourque, G., Frontini, M., Danesh, J., Roberts, D. J., Ouwehand, W. H., Butterworth, A. S., and Soranzo, N. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* **167**(5), 1415–1429.e19, 11 (2016).

313. Sirajuddin, M., Rice, L. M., and Vale, R. D. Regulation of microtubule motors by tubulin isotypes and post-translational modifications. *Nat Cell Biol* **16**(4), 335–44, Apr (2014).

314. Ikegami, K. and Setou, M. Ttll10 can perform tubulin glycylation when co-expressed with ttll8. *FEBS Lett* **583**(12), 1957–63, Jun (2009).

315. Khan, A. O., Simms, V. A., Pike, J. A., Thomas, S. G., and Morgan, N. V. Crispr-cas9 mediated labelling allows for single molecule imaging and resolution. *Sci Rep* **7**(1), 8450, Aug (2017).

316. Beghin, A., Kechkar, A., Butler, C., Levet, F., Cabillic, M., Rossier, O., Giannone, G., Galland, R., Choquet, D., and Sibarita, J.-B. Localization-based super-resolution imaging meets high-content screening. *Nat Methods* **14**(12), 1184–1190, Dec (2017).